Tue Apr

```
April 3, 2001, 11:38:17 ; Search time 16.8 Seconds (without alignments) 459.988 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                US-09-441-723-1
1185
1 MGPLPRTVELFYDVLSPYSW.....AHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    // SIDSI/gcgdata/geneseqp/AA1980.DAT:*
// SIDSI/gcgdata/geneseqp/AA1981.DAT:*
// SIDSI/gcgdata/geneseqp/AA1991.DAT:*
// SIDSI/gcgdata/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqp/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1992.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1997.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1997.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1997.DAT:*
// SIDSI/gcgdata/geneseqf/geneseqp/AA1997.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
/SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      268485 segs, 34193795 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A_Geneseq_36:*
                                                                                                                                                                                                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                          Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES	ID Description	Y77499 Novel human glutat	Y59988 Human endometrium	Y32947 Mutant threonine d		~	Y32939 Mutant threonine d		Y32951 Wild type threonin			Y32952 Mutant threonine d	0 1000000
		21	20	20	20	20	20	50	20	20	20	20	0
	Query Match Length DB	226	256	545	545	290	592	592	592	592	592	9	000
dЬ	Query	100.0	100.0	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	
	Score	1185	1185	83	89	68	88	89	68	83	8	89	ć
	Result No.		8	m	❖	ស	9	7	æ	6	10	11	,

Feedback insensiti Inosine-guanosine Mutant threonine d Feedback insensiti Mouse STE20-relate Amino acid sequenc	t thre ST re ST re ST re acid ydia p tomyce tomyce	maritima the frendoxin red man adrenodoxin red human protein rotease. Pyro prococcus furly 8855926 Seq I uman secreted ovine adrenal	Human secreted pro Human GEK2 protein Full length human TATA-binding prote Human TATA-binding TATA-binding prote Arabidopsis thalia Arabidopsis thalia
Y05704 W04266 Y32942 Y05706 Y55957 Y33825	732942 732942 732942 745707 745707 745707 745707 745707 745707 745707 745707 745707 745707 745707 745707 745707	Y56814 R27786 R27786 R30124 W24124 Y91663 P94263	Y91669 Y55939 Y55936 R56501 W06091 W30611
200 200 200 200 200	20 20 20 20 20 20 20 20 20	220 117 120 120 120	20 11 10 20 20 20
609 308 539 311 339	532 532 1073 1073 341 659 216 317	2317 4497 1398 1398 1398 142	00000000000000000000000000000000000000
			, , , , , , , , , , , , , , , , , , ,
	2008 2008 2009 2009 2009 2009 2009 2009	വിവിവ വിവിവ	74 73.5 73.5 73.5 73.5
113 114 116 117	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	0.00 H G E E E E E E E E E E E E E E E E E E	38 39 44 44 54 54

ALIGNMENTS

Isolated nucleic acid encoding glutathione S-transferase useful in the production of agents for preventing, diagnosing and treating diseases associated with cell proliferation -Glutathione S-transferase; human; GSTS; cancer; immune disorder; gene therapy; diagnosis; treatment; drug screening. Novel human glutathione S-transferase, GSTS. Corley NC; Claim 8; Fig 1A-C; 27pp; English Y77499 standard; Protein; 226 AA. Lal P, 97US-0978174. 97US-0978174. (first entry) (INCY-) INCYTE PHARM INC. Hillman JL, Shah P, WPI; 2000-205204/18. N-PSDB; 202599. 25-NOV-1997; Homo sapiens 25-NOV-1997; 05-JUN-2000 US6030809-A. 29-FEB-2000 X77499; RESULT Y77499

Uncleotide sequences encoding GSTS were initially isolated from a urologic cDNA library, and subsequently extended using cDNA libraries.

Uncleotide sequences encoding GSTS were initially isolated from a urologic cDNA library, and subsequently extended using cDNA libraries accorded by a a consensus CDNA. GSTS, and nucleotides which encode it may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GSTS expression, such as cancers and immune disorders associated with reduced expression or activity of GSTS, and in antisense therapy for disorders associated with increased GSTS expression or activity. They may also be used for the recombinant production of GSTS, and as a source of probes and primers to detect and quantitate the presence of similar nucleic acid sequences, particularly for the diagnosis of GSTS associated disorders. GSTS proteins may be used quantitate the presence of similar nucleic acid sequences, particularly for the diagnosis of GSTS antibodies against GSTS expression and activity. The anti-GSTS antibodies against GSTS expression and activity and activity. The anti-GSTS antibodies and GSTS entagonists may also be used activity. The anti-GSTS antibodies and activity. Antigonists of GSTS expression and function may be used treat immune disorders (e.g., AIDS, anamemia, asthma, Crohn's disease, irritable bowel syndrome, multiple sclerosis, osteoarthritis, microbial infections) and cancers (e.g., diagnoses). The mati-GSTS antibodies may also be used as allowed. The anti-GSTS antibodies may also be used as sequence represents a novel human glutathione S-transferase, GSTS. diagnostic agents. ****8888888888888888888888888888**

226 AA; Seguence

```
ó
                                                                                                   121 LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                 0; Gaps
                                                                 1 MGPLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLP 60
 100.0%; Score 1185; DB 21; Length 226; 100.0%; Pred. No. 4.5e-122; tive 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                     FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                  Query Match
Best Local Similarity 100.
Matches 226; Conservative
                                                                                                                                                                                                    181
                                                                         g
                                                                                                   ò
                                                                                                                           g
                                                                                                                                                  ò
                                                                                                                                                                           a
                                                                                                                                                                                                    à
```

Endometrium; human; tumour; cancer; anticancer; cytostatic; EST: treatment; uterine; gene therapy; expressed sequence tag. Human endometrium tumour EST encoded protein 48. Y59988 standard; Protein; 256 AA. 98DE-1017948. 98DE-1017948. 31-JAN-2000 (first entry) DE19817948-A1 Homo sapiens. 17-APR-1998; 1.7-APR-1998; 21-0CT-1999 RESULT Y59988

Threonine dehydratase/deaminase; TD; feedback insensitive mutant; molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.

Arabidopsis thaliana

WO9941395-A1. Synthetic.

Mutant threonine dehydratase/deaminase protein sequence.

09-NOV-1999 (first entry)

Y32947 standard; Protein; 545 AA.

```
that are highly expressed in uterine tumour tissue and which have anticancer and cytostatic activity. (A) are used (1) for recombinant expression of polypeptides (B) and (11) to isolate complete genes. (B) are used (1) to identify agents suitable for treatment of uterine or endometrial cancer; (11) directly for treating these forms of cancer (1 cluding expression from gene therapy vectors) and (11) for generation of specific antibodies. (A) are identified by assembling ESTS (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the generation contains the fact that ESTS from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of cocurrence in a particular tissue. Y59941-Y60128 represent protein fragments encoded by the human endometrium tumour cDNA library derived EST fragments represented in 241981-242121.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                This invention describes novel human nucleic acid (cDNA) sequences (A),
                                             Dahl E;
                                                                                                                                                  New nucleic acid sequences expressed in uterine cancer tissues, and derived polypeptides, for treatment of uterine and endometrial cancer and identification of therapeutic agents \, -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LWWRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        151 lwmrvwsrneditepgsilaaaekagmsaegaggllekiatpkvknglketteaacryga 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MGPLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Score 1185; DB 20; Length 256; 100.0%; Pred. No. 5.4e-122; ive 0; Mismatches 0; Indels 0;
                                         Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          211 fglþitvahvdggthmlfgsdrmellahllgekwmgpippavnarl 256
(META-) METAGEN GES GENOMFORSCHUNG MBH.
                                                                                                                                                                                                                                     Claim 23; Page 294; 444pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match
Best Local Similarity 100.0
Matches 226; Conservative
                                                                                    WPI; 1999-591957/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 256 AA;
                                                                                                         N-PSDB; Z41996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X32947
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
```

```
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                            This sequence represents a mutant Arabidopsis thaliana threonine dehydratased/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence iss used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of synthesis of the degradable biopolymer poly(hydroxybutyrate). Also TD-expressing plants permit use of the isoleucine structural analogues as markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105 -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113 gvicssagnhaggvalsasklgctavivmpvttpeikwqavenl------- 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 ppklplpr-----lkvspnslq----ypagylgavpertneaengsiaeameyltni 52
                                                                                                                                                                           New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Threonine dehydratase/deaminase; OMR1 gene; feedback inhibition; transgenic plant; selectable marker; isoleucine; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                80;
                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 20; Length 545;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Feedback insensitive threonine dehydratase/deaminase.
                                                                                                                        Mourad GS, Pareddy DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          157 ---gatvvlfgdsydqaqahakiraeeegltfipp 188
                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 7.5%; Score 89; DB Best Local Similarity 22.8%; Pred. No. 0.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              29; Mismatches
                                                                                                                                                                                                                        Claim 13; Page 131-134; 194pp; English,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Y05711 standard; Protein; 545 AA
                                                                                      ) DOW AGROSCIENCES LLC. ) PURDUE RES FOUND.
                                                       98WO-US14362.
98US-0074875.
                                99WO-US00560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-JUL-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              49; Conservative
                                                                                                                      Larrinua IM, Merlo DJ,
                                                                                                                                                                                                                                                                                                                                                                                  should be safe to use.
                                                                                                                                          WPI; 1999-527375/44.
                                                                                                                                                                                                                                                                                                                                                                                                        545 AA;
                                                                                                                                                        N-PSDB; Z11205
                                08-JAN-1999;
                                                                 17-FEB-1998;
                                                                                       (DOWC ) DOW
          19-AUG-1999
                                                       10-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Y05711;
                                                                                                 PURD
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
Y05711
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
The present sequence represents an Arabidopsis thallana mutant threonine dehydratass/deaminase (TD) protein which, unlike threonine dehydratass/deaminase (TD) protein which, unlike wild-type TD, is insensitive to feedback inhibition by isoleucine. Claimed polynucleotides (see X25332-40), originally isolated and cloned from A. thallana mutated line GMID (omr1/omr1), encode feedback insensitive TD that can be used to transform a wide variety of plants, fund; batch and yeast. Mutant TD differs from the wild type enzyme only by an R499C amino acid substitution regulatory region R4, and by an R544H substitution in regulatory in regulatory region R6. Mutant TD is not only insensitive to feedback inhibition by isoleucine, but is also insensitive to structural analogues of isoleucine that are toxic to plants and microorganisms which synthesise only wild type TD. Nucleotide sequences encoding mutated forms of TD can therefore be used to create cells that are microorganisms of incontinuous and interest forms of TD can therefore be used to create cells that are
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             53 lstkvydiaiesplqlakklskrlgvrmylkredlqpvfsfklrgaynmmvklpadqlak 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----POSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences encode mutant threonine dehydratase/deaminase - which is insensitive to feedback inhibition, useful as a selective marker to produce transformed cells resistant to toxic isoleucine analogues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               insensitive to compounds normally toxic to cells expressing only wild type TD enzymes, and thus may be used to provide a biochemical selectable marker. Transformants that express the mutant TD show increased levels of isoleucine production, and thus provide an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.5%; Score 89; DB 20; Length 545; 22.8%; Pred. No. 0.3; ative 29; Mismatches 57; Indels E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   105 ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE----
                                                                                                                                                                    "regulatory region R4"
                                                                                                                                                                                                                                   /note= "regulatory region R6"
                                                                                                                                                                                                                                                                                              'note= "Arg in wild-type TD"
                                                                                                                                                                                                                                                                                                                                                                /note= "Arg in wild-type TD"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 71-73; 120pp; English.
                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US14362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0074875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0052096
                                                                                                                                                                    /note= "r
489..507
                                                                                                                                 ..457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (PURD ) PURDUE RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       improved nutrient source.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
Arabidopsis thaliana
                                                                                                                                                                                                                                                               Misc-difference 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-120860/10.
                                                                                                                                                                                                                                                                                                                             Misc-difference 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      545 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; X25340.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9902656-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mourad GS;
                               Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seguence
                                                                                                                                 Region
                                                                                                                                                                                                   Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
```

11;

Tue Apr

ď q

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This sequence represents a mutant Arabidopsis thaliana threonine dehydratase/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of intermediates in biosynthesis of isoleucine (e.g. 2-oxobutyrate, for synthesis of the degradable biopolymer poly(hydroxybutyrate)). Also TD-expressing plants permit use of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antibiotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                                                                                      Threonine dehydratase/deaminase; TD; feedback insensitive mutant; molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.5%; Score 89; DB 20; Length 590;
22.8%; Pred. No. 0.34;
Live 29; Mismatches 57; Indels 80; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                         Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pareddy DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE---
189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                     Example 3; Page 119-123; 194pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mourad GS,
                                                                                                                       Y32950 standard; Protein; 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (DOWC ) DOW AGROSCIENCES LLC (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US14362.
98US-0074875.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-US00560.
                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Larrinua IM, Merlo DJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             should be safe to use.
                                                                                                                                                                                                                                                                                                                                                 Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-527375/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             590 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; Z11208.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUL-1998;
17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                   WO9941395-A1
                                                                                                                                                                                        09-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                      19-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                          Y32950;
                                                                                     Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                   RESULT
                                                                                                         X32950
                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This sequence represents a mutant Arabidopsis thaliana threonine dehydratase/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of intermediates in biosynthesis of isoleucine (e.g. 2-oxobutyrate, for synthesis of the degradable bioplymer poly(hydroxybutyrate)). Also TD-expressing plants permit use of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antibiotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                                                                                                                                                                                                                                                      molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
--LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                      158 gvicssagnhaggvalsasklgctavivmpvttpeikwgavenl---------201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell
                                                                                                                                                                                                                                                                                                                                                                                                         Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 7.5%; Score 89; DB 20; Length 592; Best Local Similarity 22.8%; Pred. No. 0.34; Matches 49; Conservative 29; Mismatches 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                                                                                                                                                                        Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Merlo DJ, Mourad GS, Pareddy DR;
                                                                                                                                189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                 202 ---gatvvlfgdsydqaqahakiraeeegltfipp 233
                                                                                                                                                                                                                                                                 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 13; Fig 8; 194pp; English.
                                                                                                                                                                                                                                                               Y32939 Standard; Protein; 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (DOWC ) DOW AGROSCIENCES LLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0074875.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-US00560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US14362
                                                                                                                                                                                                                                                                                                                                        09-NOV-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    should be safe to use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-527375/44.
                                                           142 AEKAGMSAEQAQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         592 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; Z11197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Carrinua IM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9941395-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                   Y32939;
                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                             Y32939
                                                                                                                                                                                                                                                                                                       ò
                                                                                               g
                                                                                                                                    ò
                                                                                                                                                                     g
```

DB 20; Length 592;

592 AA;

Sequence

S

S

Tue Apr

```
This sequence represents a mutant Arabidopsis thaliana threonine dehydratass/deaminase (TD) protein of the invention. The protein is a feedback linsensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of synthesis of isoleucine (and thus better nutritional value) or of synthesis of the degradable biopolymer poly(hydroxybutyrate). Also TD-expressing plants permit use of the isoleucine structural analogues as herbiddes. The DNA sequences are alternatives for antibiotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   molecular marker; isoleucine toxic structural analog registance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                             ------PQSILAA 141
                                                                                                                                               ------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
 ---LEKGSLS-AMRFLT-- 104
                                  52 ppklplpr-----lkvspnslq----ypagylgavpertneaengsiaeameyltni 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                   160 gvicssagnhaggvalsasklgctavivmpvttpeikwgavenl----
                                                                        105 ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pareddy
                                                                                                                                                                                                                                            204 ---gatvvlfgdsydqaqahakiraeeegltfipp 235
 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-
                                                                                                                                                                                                                     189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 3; Page 126-129; 194pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mourad GS,
                                                                                                                                                                                                                                                                                                                                                     A
                                                                                                                                                                                                                                                                                                                                                   Y32948 standard; Protein; 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOWC ) DOW AGROSCIENCES LLC. PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US14362.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99WO-US00560
                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       carrinua IM, Merlo DJ,
                                                                                                                                             142 AEKAGMSAEQAQG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-527375/44.
N-PSDB; 211206.
                                                                                                                                                                                                                                                                                                                                                                                                                          09-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9941395-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .0-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                       X32948;
                                                                                                       g
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                g
                                                                                                                                             ò
                                                                                                                                                                                                                     ò
                                                                        ò
```

should be safe to use.

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  dehydratase/deaminase (TD) protein. The invention relates to mutants of the encoded protein, that are feedback insensitive TD mutants. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of intermediates in blosynthesis of isoleucine (e.g. 2-oxobutyrate, for synthesis of the degradable blopolymer poly(hydroxybutyrate)). Also TD-expressing plants permit use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker.
                                                                                                                                                           -LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                  55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                       -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE------PQSILAA 141
                                  Gaps
                                                                                                     52 ppklplpr-----lkvspnslq----ypagylgavpertneaengslaeameyltni 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sequence is the wild type Arabidopsis thaliana threonine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wild type threonine dehydratase/deaminase protein sequence.
                                  57; Indels
                                                                                                                                                                                                                                              160 gvicssagnhaggvalsasklgctavivmpvttpeikwgavenl---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pareddy
                  0.34;
                                                                                                                                                                                                                                                                               189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                          204 ---gatvvlfgdsydqaqahakiraeeegltfipp 235
7.5%; Score 89; DB 22.8%; Pred. No. 0.34 cive 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Page 86-89; 194pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mourad GS,
                                                                                                                                                                                                                                                                                                                                                                                                        A
                                                                                                                                                                                                                                                                                                                                                                                                      Y32951 standard; Protein; 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (DOWC ) DOW AGROSCIENCES LLC (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99WO-US00560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US14362
98US-0074875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Merlo DJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1999-527375/44.
 Query Match
Best Local Similarity
Matches 49; Conserv
                                                                                                                                                                                                              142 AEKAGMSAEQAQG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; Z11209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Larrinua IM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09941395-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUL-1998;
17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                        X32951;
                                                                                                                                         105
                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                       732951
                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                         ò
                                                                                                                                                                         qq
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                             q
```

ဖ

```
11;
of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antiblotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it should be safe to use.
                                                                                                                                                                                                                                        PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                           ---PQSILAA 141
                                                                                                                                                                                                                                                                              142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                   52 ppklplpr-----lkvspnslq----ypagylgavpertneaengsiaeameyltni 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New sequences encode mutant threonine dehydratase/deaminase - which is insensitive to feedback inhibition, useful as a selective marker to produce transformed cells resistant to toxic isoleucine analogues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Threonine dehydratase/deaminase; OMR1 gene; feedback inhibition; transgenic plant; selectable marker; isoleucine.
                                                                                                                                          80;
                                                                                                               Length 592;
                                                                                                            7.5%; Score 89; DB 20; Length 59:
22.8%; Pred. No. 0.34;
.ive 29; Mismatches 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Arabidopsis wild-type threonine dehydratase/deaminase.
                                                                                                                                                                                                                                                                                                  105 ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "regulatory region R4" 536..554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note- "regulatory region R6"
                                                                                                                                                                                                                                                                                                                                                        204 ---gatvvlfgdsydqaqahakiraeeegltfipp 235
                                                                                                                                                                                                                                                                                                                                     189 HVDGQTHMLFGSDRMELLAHL-~-LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "transit peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91.592
/note= "mature protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
1..90
                                                                                                                                                                                                                                                                                                                                                                                                                                    x05702 standard; Protein; 592 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0074875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97us-0052096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-JUL-1999 (first entry)
                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           486..504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-120860/10.
                                                                                                             Query Match
Best Local Similarity
Matches 49; Conserv
                                                                    592 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; X25331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9902656-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-JAN-1999
                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Y05702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Region
                                                                                                                                                                   22
$$$$$$$
                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
The present sequence represents Arabidopsis thaliana var. Columbia wild-type threonine dehydratase/deaminase (TD), the first enzyme of the isoleucine biosynthetic pathway. TD is encoded by the OMRI gene (see x25331) of A. thaliana. The invention provides nucleotide sequences (see x25332-40), originally isolated and cloned from A. thaliana mutated line GMIIb (omri/Omri), which encode feedback insensitive TD that can be used to transform a wide variety of plants, fungi, bacteria and yeast. The mutated form of TD differs from the wild type only by 2 point mutations (C to T at nucleotide 1495, and G to A at nucleotide 1531), which result in an R499C amino acid substitution in the regulatory region R4 of TD, and an R544H substitution in regulatory region R6. These forms of TD are not only insensitive to structural analogues of isoleucine that are toxic to plants and microorganisms which synthesise only wild-type TD. Uncleotide sequences encoding mutated forms of TD can therefore be used to create cells that are insensitive to compounds normally toxic to cells expressing only wild-type TD enzymes, and thus may be abloched as blochedical selectable marker. Transformants
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       isoleucine production, and thus provide an improved nutrient source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Threonine dehydratase/deaminase; OMR1 gene; feedback inhibition; transgenic plant; selectable marker; isoleucine; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                            harboring a nucleotide sequence comprising a promoter operably linked to a mutated TD sequence demonstrate increased levels of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Feedback insensitive threonine dehydratase/deaminase precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.5%; Score 89; DB 20; Length 592;
22.8%; Pred. No. 0.34;
Live 29; Mismatches 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               160 gvicssagnhaggvalsasklgctavivmpvttpeikwgavenl--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          204 ---gatvvlfgdsydgagahakiraeeegltfipp 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note- "transit peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "mature protein"
                  Disclosure; Page 50-53; 120pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Y05703 standard; Protein; 592 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-JUL-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  486. 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arabidopsis thallana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            592 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Sim-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X05703;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Y05703
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
```

/note= "regulatory region R4"

11;

us-09-441-723-1.rag

```
threonine dehydratase/deaminase (TD) precursor which, unlike wild-type TD, is insensitive to feedback inhibition by isoleucine. Claimed polynucleotides (see X25332-40), originally isolated and claimed polynucleotides (see X25332-40), originally isolated and claimed polynucleotides (see X25332-40), originally isolated and feedback insensitive TD that can be used to transform a wide feedback insensitive TD that can be used to transform a wide variety of plants, fungil, bacteria and yeast. Mutant TD differs from the wild-type enzyme only by an R499C amino acid substitution regulatory region R4, and by an R544H substitution in regulatory in regulatory region R4, and by an R544H substitution in regulatory in region R6. Mutant TD is not only insensitive to feedback inhibition by isoleucine, but is also insensitive to structural analogues of isoleucine that are toxic to plants and microorganisms which synthesise only wild-type TD. Nucleotide sequences encoding insensitive to compounds normally toxic to cells expressing only wild-type TD enzymes, and thus may be used to provide a blochemical energy of increased levels of isoleucine production, and thus provide an
                                                                                                                                                                                                                                                                                                                                                                                                                                                New sequences encode mutant threonine dehydratase/deaminase - which is insensitive to feedback inhibition, useful as a selective marker to produce transformed cells resistant to toxic isoleucine analogues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present sequence represents an Arabidopsis thaliana mutant
536..554
/note= "regulatory region R6"
                                                       'note- "Arg in wild-type TD"
                                                                                              /note= "Arg in wild-type TD"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 53-56; 120pp; English
                                                                                                                                                                                                                                                       98US-0074875.
                                                                                                                                                                                                                  98WO-US14362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                improved nutrient source.
                                                                                                                                                                                                                                                                                                                (PURD ) PURDUE RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-120860/10.
N-PSDB; X25332.
                                       Misc-difference 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     592 AA;
                                                                            Misc-difference
                                                                                                                                                                                                                                                                           10-JUL-1997;
                                                                                                                                                                                                                10-JUL-1998;
                                                                                                                                     WO9902656-A1
                                                                                                                                                                             21-JAN-1999
                                                                                                                                                                                                                                                                                                                                                     Mourad GS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
  Region
```

11; -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE------PQSILAA 141 ------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188 55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104 Gaps 52 ppklplpr-----lkvspnslq----ypagylgavpertneaengsiaeameyltni 99 80; DB 20; Length 592; 57; Indels gvicssagnhaggvalsasklgctavivmpvttpeikwgavenl--HVDGQTHMLFGSDRMELLAHL --- LGEKWMGPIPP 220 7.5%; Score 89; DB 22.8%; Pred. No. 0.34. iive 29; Mismatches Best Local Similarity 22.8 Matches 49; Conservative 142 AEKAGMSAEQAOG-Query Match 105

q

ò

q .8 g

ò

11;

142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188 gvicssagnhaqgvalsasklgctavivmpvttpeikwqavenl------ 211

168

105 -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAA 141

ð QQ ò q

```
dehydratase/deaminase (TD) protein of the invention. The protein is dehydratase/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of intermediates in biosynthesis of isoleucine (e.g. Coxobutyrate, for synthesis of the degradable biopolymer poly(hydroxybutyrate)). Also TD-expressing plants permit use of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antiblotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                              molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.5%; Score 89; DB 20; Length 600;
22.8%; Pred. No. 0.35;
tive 29; Mismatches 57; Indels 80; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                            Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This sequence represents a mutant Arabidopsis thaliana threonine
                                                                                                                                            Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pareddy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mourad GS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 2; Fig 5; 194pp; English
                                    Y32952 standard; Protein; 600 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (DOWC ) DOW AGROSCIENCES LLC (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                 99WO-US00560.
                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-US14362.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0074875.
                                                                                                          09-NOV-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 22.8
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Merlo DJ,
                                                                                                                                                                                                                                                                        Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1999-527375/44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     600 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; Z11212.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Larrinua IM,
                                                                                                                                                                                                                                                                                                                           WO9941395-A1
                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                               .9-AUG-1999
                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                      Y32952;
 Ξ
RESULT
Y32952
                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
```

```
This sequence represents a mutant Arabidopsis thaliana threonine dehydratase/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transformants with increased levels of isoleucine (and thus better nutritional value) or of intermediates in biosynthesis of isoleucine (e.g. 2-oxobutyrate, for synthesis of the degradable bioplymer poly(hydroxybutyrate)). Also TD-expressing plants permit use of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antibiotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it
                                                                                                                                                                                                                                                                                                molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New nucleic acid encoding threonine dehydratase deaminase resistant to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                                                                                                                               Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 609;
                                                                                                                                                                                                                                           Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pareddy DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 7.5%; Score 89; DB 20; Best Local Similarity 22.8%; Pred. No. 0.36; Matches 49; Conservative 29; Mismatches 57;
   189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                      Larrinua IM, Merlo DJ, Mourad GS,
                                                                                                                                Y32940 standard; Protein; 609 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 13; Fig 6b; 194pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (DOWC ) DOW AGROSCIENCES LLC. (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US14362.
98US-0074875.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99WO-US00560
                                                                                                                                                                                                      09-NOV-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          should be safe to use.
                                                                                                                                                                                                                                                                                                                                                                         Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-527375/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             609 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; Z11198.
                                                                                                                                                                                                                                                                                                                                                                                                                              WO9941395-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-AUG-1999.
                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                             RESULT 12
                                                                                                             X32940
ò
```

```
-- LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New sequences encode mutant threonine dehydratase/deaminase - which is insensitive to feedback inhibition, useful as a selective marker to produce transformed cells resistant to toxic isoleucine analogues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  threonine dehydrates/deminate (TD) which, unlike the wild-type enzyme (see Y05702), is insensitive to feedback inhibition by isoleucine. It is encoded by the mutant omri gene (see X25333) of A. thaliana mutated line GMID (omr1/comr1). This is one of 9 claimed polynucleotides (see X2533-40), originally isolated and cloned from line GMID, which encode feedback insensitive TD that can be used to transform a wide variety of plants, fund; bacteria and yeast. These polynucleotides differ from the wild-type only by 2 point mutations, which result in R499C and R554H amino acid substitutions (numbering according to wild-type TD) in the R4 and R6 regulatory regions. These forms of TD are not only insensitive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R6 regulatory regions. These forms of TD are not only insensitive to feedback inhibition by isoleucine, but are also insensitive to
                                                                                                                                                                                                                                                                                                                                                                     Threonine dehydratase/deaminase; omr1 gene; feedback inhibition;
transgenic plant; selectable marker; isoleucine; mutant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence represents an Arabidopsis thaliana mutant
                                                                                                                                                                                                                                                                                                                                    Feedback insensitive threonine dehydratase/deaminase.
                                                                               177 gvicssagnhaqgvalsasklgctavivmpvttpeikwgavenl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Arg in wild-type enzyme"
503..521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    note= "Arg in wild-type enzyme"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..521
ste= "regulatory region R4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="regulatory region R4" 553..572 /note="regulatory region R6"
                                                                                                                189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                    221 ---gatvvlfgdsydqaqahakiraeeegltfipp 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 56-59; 120pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                 Y05704 standard; Protein; 609 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0074875.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US14362
                                                                                                                                                                                                                                                                                                   19-JUL-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (PURD ) PURDUE RES FOUND.
                                                                                                                                                                                                                                                                                                                                                                                                                      Arabidopsis thallana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-120860/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 516
                                              142 AEKAGMSAEQAQG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; X25333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                  Y05704;
                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                  X05704
                                                                                                                                                                                                                                                                 g
                                                                                                                  ò
                                                                                                                                              qq
```

11;

Gaps

57; Indels 80;

55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104

ga à

105 -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE--------PQSILAA 141

Tue Apr

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This sequence represents the inosine-guanosine kinase sequence derived from Exiguobacterium sp. This sequence was used in the production of
                                                                                                                                                                                                                                                                                                                                                69 ppklplpr-----lkvspnslq----ypagylgavpertneaengsiaeameyltni 116
                                                                                                                                                                                                                                                                                                                                  -----PQSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                           142 AEKAGMSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                  55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                           Gaps
structural analogues of isoleucine that are toxic to plants and microorganisms which synthesise only wild-type TD. Nucleotide sequences encoding mutated forms of TD can therefore be used to create cells that are insensitive to compounds normally toxic to cells expressing only wild-type TD enzymes, and thus may be used to provide a biochemical selectable marker. Transformants to provide a biochemical selectable marker. Transformants linked to a mutated TD sequence comprising a promoter operably linked to a mutated TD sequence demonstrate increased levels of isoleucine production, and thus provide an improved nutrient source.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  for use
                                                                                                                                                                                                                                        80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Inosine-guanosine kinase; 5'-inosinic acid; 5'-guanylic acid; Corynebacterium ammoniagenes; ATP; seasoning; food.
                                                                                                                                                                                                       DB 20; Length 609;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Prodn. of 5'-inosinic acid or 5'-guanylic acid - using C. ammoniagenes transformed by inosine-guanosine kinase gene,
                                                                                                                                                                                                                                      57; Indels
                                                                                                                                                                                                                                                                                                                              ----AVNLEHP-EMLEKASRELWMRVWSRNEDITE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Utagawa T;
                                                                                                                                                                                                                      Pred. No. 0.36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       189 HVDGQTHMLFGSDRMELLAHL --- LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---gatvvlfgdsydqaqahakiraeeegltf1pp 252
                                                                                                                                                                                                                                      29; Mismatches
                                                                                                                                                                                                       7.5%; Score 89; 22.8%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 14; Page 65-66; 72pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Usuda Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Exiguobacterium sp. (ATCC 35652)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W04266 standard; Protein; 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96WO-JP00761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95JP-0177900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95JP-0102888
                                                                                                                                                                                                    Query Match 7.5%
Best Local Similarity 22.8%
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Inosine-guanosine kinase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (AJIN ) AJINOMOTO CO INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Shimaoka M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1996-455349/45.
N-PSDB; T33972.
                                                                                                                                                           609 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             food seasoning
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22-MAR-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kawasaki H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24-MAR-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JUN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        03-OCT-1996
                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W04266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
```

```
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                dehydratase/deaminase (TD) protein of the invention. The protein is a feedback insensitive mutant. The TD DNA sequence is used as molecular marker (imparting resistance to toxic structural analogues of isoleucine) for selecting transformed cells and to produce transforments with increased levels of isoleucine (and thus better nutritional value) or of
5'-inosinic acid or 5'-guanylic acid by reacting an organism, pref. Corynebacterium ammoniagenes, capable of regenerating ATP, containing this recombinant DNA, with inosine, guanosine or their precursors, and an energy and phosphoric acid source. 5'-inosinic acid and 5'-guanylic acid are useful as seasonings for food.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      molecular marker; isoleucine toxic structural analog resistance; isoleucine production; biosynthesis; degradable biopolymer; herbicide; polyhydroxybutyrate; antibiotic resistance marker; mutein.
                                                                                                                                                                                                        63 GLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRELW 122
                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                              New nucleic acid encoding threonine dehydratase deaminase resistant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Threonine dehydratase/deaminase; TD; feedback insensitive mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This sequence represents a mutant Arabidopsis thaliana threonine
                                                                                                                                                                              33;
                                                                                                                                               Length 308;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       to feedback inhibition, useful as selection marker for cell transformation and to impart herbicide resistance
                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mutant threonine dehydratase/deaminase protein sequence.
                                                                                                                                                                                                                                                                     123 MRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVK----
                                                                                                                                                                           57;
                                                                                                                                             7.0%; Score 83.5; DB 17; 25.8%; Pred. No. 0.52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pareddy DR;
                                                                                                                                                                                                                                                                                                                                                        171 TTEAACR----YGAFGLPITVAHVDGQTHMLFG 199
                                                                                                                                                                           25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 13; Page 106-109; 194pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mourad GS,
                                                                                                                                                                                                                                                                                                                                                                                                                                             Y32942 standard; Protein; 539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (DOWC ) DOW AGROSCIENCES LLC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-US00560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US14362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0074875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-NOV-1999 (first entry)
                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (PURD ) PURDUE RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Merlo DJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-527375/44.
                                                                                                                                                          Local Similarity
nes 40; Conserv
                                                                                               308 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; Z11200.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Larrinua IM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9941395-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-FEB-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-AUG-1999,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                 Sequence
                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
Y32942
                                                                                                                                                              Best Loc
Matches
    888888×8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                              Op
```

```
10;
intermediates in biosynthesis of isoleucine (e.g. 2-oxobutyrate, for synthesis of the degradable biopolymer poly(hydroxybutyrate)). Also TD-expressing plants permit use of the isoleucine structural analogues as herbicides. The DNA sequences are alternatives for antibiotic resistance markers (which are potentially harmful to the environment). Since no human analog of TD exists (humans can not synthesize isoleucine), it should be safe to use.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               105 -AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAAAEKAG 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : | | :: :| |: | :|:: : ||:
52 ydiaiesplqlakklskrlgvrmylkredlqpvfsfklrgaynmmvklpadqlakgvics 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  147 MSAEQAQG------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVDGQ 193
                                                                                                                                                                                                                                                                                                                                                                   59 LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT----- 104
                                                                                                                                                                                                                                                                                                                     79; Gaps
                                                                                                                                                                                                                                                                                                                                                                                             Query Match 7.0%; Score 83.5; DB 20; Length 539; Best Local Similarity 21.9%; Pred. No. 1.2; Matches 46; Conservative 29; Mismatches 56; Indels 79.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               112 sagnhaqgvalsasklgctavivmpvttpeikwqavenl-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             194 THMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            539 AA;
                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
       8×949999988
```

Search completed: April 3, 2001, 11:39:01 Job time: 44 sec

```
CLASSIFICATION:
PRIOR APPLICATION
APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 226 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: Herewith
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT04
; CLONE: 1554593
US-08-978-174-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 94304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1
US-08-978-174-1
                          69
69
69
68
68
68
68
68
68
67
67
67
67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 9, Appli
Sequence 3, Appli
Sequence 32, Appli
Sequence 32, Appli
Sequence 32, Appli
Sequence 6, Appli
                                                                                                                                                             (without alignments)
276.074 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl
Appl
Appl
Appl
Appl
Appl
Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                   1185
1 MGPLPRTVELFYDVLSPYSW......AHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 14,
Sequence 14,
Sequence 14,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1
Sequence 2
Sequence 3
Sequence 1
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5
Sequence 1
                                                                                                                                     3, 2001, 11:38:17 ; Search time 14.7 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1: /cgn2_6/ptcdatt/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptcdatts/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptcdatts/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptcdatts/2/iaa/pcTUS_COMB.pep:*
5: /cgn2_6/ptcdatts/2/iaa/pcTUS_COMB.pep:*
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-188-582-32
US-08-646-715-32
US-08-473-553A-6
US-08-473-553A-2
US-08-698-551-14
                                                                                                                                                                                                                                                                                                                                                                                                                                         fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US95-12724-14
US-08-922-170B-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-05008-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-533-901B-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-839-032A-14
US-08-839-031A-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-07-955-905A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-484-101B-38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-198-446B-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-07-857-224B-51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-855-714-3
US-09-066-075-2
US-08-518-615A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-564-090A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PCT-US94-06698-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ·08-473-553A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-602-228-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-870-693-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-951-889-2
US-08-075-193-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -08-750-532-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-978-174-1
                                                                                                                                                                                                                                                                                                                                                                                             174772 segs, 17957048 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                         protein search, using sw model
                                                                                                                                                                                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length: 0
length: 2000000000
                                                                                                                                                                                                                               US-09-441-723-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Match Length
                                                                                                                                     April
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88087

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007

8007
                                                                                                                                                                                                                                                   Perfect score:
                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                       OM protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum DB Maximum DB
                                                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database
                                                                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
```

Length 226;

DB 3;

100.0%; Score 1185;

```
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                Sednence Sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1
                                                                   Sequence
                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/08978174

Patent No. 6030809

GENERAL INFORMATION:
APPLICANT: Shah, Purvi
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: New GLUTATHIONE-S-TRANSFERASE
NUMBER OF SEQUENCES: 3
US-07-623-033-2
US-08-059-031-5
US-08-059-035-5
PCT-US91-07290-5
US-08-758-280-1
US-08-758-280-2
US-08-964-614A-1
US-08-964-614A-1
US-08-964-614A-2
US-08-973-831-2
US-09-354-129-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
COMPUTER: IBM COMPATIBLE
SOFTWARE: FASTER: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-149-103-4
US-08-451-883-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Incyte Pharmaceuticals, Inc
STREET: 3174 Porter Drive
CITY: Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                   689
9333
9333
9333
9333
933
946
846
846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS
```

ą

à

g

ò

g

ò

qq

ò

```
61 RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                       121 LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 113; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLPRIVELFYDVLS---PYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLL 59
                                                                   1 MGPLPRIVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLP
                                                                                                                                                                                                                                                                                                                                                                                                                    181 FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 FGLPTTVAHVDGKTYMLFGSDRMELLAYLLGEKWMGPVPPTLNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 3, Application US/08855714
Patent No. 5939075
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Houng, Hoo-Shu H.
APPLICANT: Warren, Richard L.
TITLE OF INVENTION: MUTANTS OF BRUCELLA MELITENSIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: John Moran, Esq.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE: John Moran, Esq.
STREET: HQ USAMRDC, Dept. of Army, Fort Detrick
CITY: Frederick
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 6.8%; Score 80.5; DB Best Local Similarity 21.7%; Pred. No. 0.59; Matches 65; Conservative 25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
CLASSIFICATION 244
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/334,129
FILING DATE: 04-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: 08/143,692
TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: BRUCELLA MELITENSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEO ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 339 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21702-5012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGINAL SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-855-714-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              å
                                                                                       g
                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                              ö
                                                                                                                                                                                                                                                            61 RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                                                                                                                      121 LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                           RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                                                                                           LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                0; Gaps
                                                                                              1 MGPLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31; Indels
                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-978-174-3

Sequence 3, Application US/08978174

Sequence 3, Application US/08978174

GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Shah, Purvi
APPLICANT: Hillman, Jennifer L.
APPLICANT: Carla, Preeti
APPLICANT: Carla, Neati
APPLICANT: Corlant Corlan
100.0%; Pred. No. 4.3e-126; iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 73.8%; Score 874; DB 3; Best Local Similarity 69.5%; Pred. No. 6.2e-91; Matches 157; Conservative 38; Mismatches 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER: IBM COMPALIDLE
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPELICATION DATA:
APPLICATION NUMBER: US/08/978,174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0430 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEC ID NO: 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               226 amino acids
                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Herewith
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIBRARY: GenBank
CLONE: ?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET: 3174 FUL
.Best Local Similarity
Matches 226; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: line
IMMEDIATE SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-978-174-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE:
                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                              61
```

14;

Tue Apr 3 11:51:22 2001

	Qy 117 ASRELWMRVWSRNEDITEPQS137
163 KH	Db 49 HVRIPIRMSTHAYAFPPYKIMDRFFKRVDEVINGALKRGLAVAINIHHYEELMN 102
Ov 138TI.AAAEK	OY 76 HLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEK 116
DD 103 DPEEHKERFLA	Matches 43; conservative 22; Mismatches 50; indeis 69; Ga
Oy 117 ASRE	.5; 50. Indole 80. Gane
Db 49 HVRIPIRWSTH	
нготопи 26 нготопи	US-09-066-075-2
Matches 43; Cons	
Query Match	TYPE: AMINO
	ŊΞ
US-08-518-615A-2	H
MOLECULE TYPE:	
<u> </u>	TELECOMMUNICATION INFORMATIC
; LENGTH: 317 A	; RECISTRATION NUMBER: 36,134 ; REFERENCE/DOCKET NUMBER: 331400-20
SEQUENCE CHARACT	
; TELEFAX: 201-	
TELECOMMONICALLO	; APPLICATION NUMBER: US/08/518,615
; REFERENCE/DOCK	
REGISTRATION	
; ATTORNEY/AGENT I	CURRENT APPLICATION DA
; APPLICATION NUI	
; PRIOR APPLICATIO	
FILLING DATE:	
CURRENT APPLICAT PAPPLICATION NO	
SOFTWARE: WOR	
COMPUTER: IBM	STREET: 6 BECKER FARM ROAD
, MEDIUM TYPE:	
ZIP: 07068	CORRESPONDENCE ADDRESS:
STATE: NEW JE	TITE OF INVENTION:
CITY: ROSELAN	APPLICANT: Mathur, E., et al.
ADDRESSEE: CEC	Patent No. GENERAL IN
ADDRESSEE: CA	: Sequence 2, Application US/09066075
NUMBER OF SEQUEN	RESULT 4 US-09-066-075-2
TITLE OF INVENTI	
GENERAL INFORMATION	Db 262 AACAISQFEQHIRAVAGLPLGNTDRHSDCVMENLIGDDIEKVPAILCEKNAVLHLYGKK 320
; Sequence 2, Applica; ; Patent No. 5962258	Qy
US-08-518-615A-2	
RESULT 5	Db 203 VPAAISVQTAEAARTAAEKLLHALDYVGVLGLEFFVLKDGTLLANEFAPRVHNS-GHWTE 261
7	Oy 147MSAEQAQGLLEKIA
700	Db 158 PAILEGEVEFEREVSVIAARDRSGNVAIFDLAENVHKDGILATST 202
Ov 194 THMLFGSDRME	VY
Db 163 KHTIIIGTAEW	
Qy 138ILAAAEK	
DD 103 DPEEHKEKFLAI	Qy 60 PRK-GLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEH 110
,	Db 47 PLPTSSPMNLKMCQSAPPTSWLKRRLFCPRPPHWKSLRTASQKSSFSTKAALK 99

```
10;
                                                                                                                                                                                                                                  IUr, E., et al.
ION: Carboxymethyl Cellulase from Thermotoga Maritima
ACES: 4
ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.5%; Score 76.5; DB 2; Length 317;
ty 21.1%; Pred. No. 1.5;
ervative 22; Mismatches 50; Indels 89; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---FP-----KDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEK 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AGMSAEQAQGILEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVDGO 193
                                                           LWMRVWSRNED-----ITEPQS------137
|| :: | :|
LWKQIADRYKDYPETLFFEILNEPHGNLTPEKWNELLEEALKVIRSIDK 162
                                       AGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVDGQ 193
                                                                                                                                                                                                                                                                                                     RELLA, BYRNE, BAIN, GILFILLAN,
CCHI, STEWART & OLSTEIN
KER FARM ROAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTEM: MS-DOS
TTEM: MS-DOS
TTEM: DERFECT 5.1
TION DATA:
TUTION DATA:
August 23, 1995
...
TUTION DATA:
TUTION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION:

AO, GREGORY D.

NUMBER: 36,134

KET NUMBER: 331400-20

CON INFORMATION:

201-994-1704

EQ 10 NO: 2:

TERISTICS:

AMINO ACIDS
                                                                                                                                                                                                                                                                                                                                                                                                              E FORM:
3.5 INCH DISKETTE
1 PS/2
                                                                                                 LLAHLLGEKWMGP 217
                                                                                                                     || || || GSEKWLGRKWGSP 225
                                                                                                                                                                                                        tion US/08518615A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VEAR
PROTEIN
                                                                                                                                                                                                                                                                                                                                                          D
RSEY
```

```
Sequence 4, Application US/08075193
Patent No. 5547868
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 415-857-0663
TELEX: 380816 COLEY PA
INFORMATION FOR SEQ ID NO: 4
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 415-494-7622
415-857-0663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               497 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 25.4%
Matches 44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                   STREET: FIVE PALO POLO POLO PALO PALO PALO ALTO STATE: CALIFORNIA COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                              94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-564-090A-4
    US-08-075-193-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-075-193-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76 HLQIPIH-----FP-----KDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEK 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    103 DPEEHKERFLALWKQIADRYKDYPETLFFEILNEPHGNLTPEKWNELLEEALKVIRSIDK 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                138 ----ILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVDGQ 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             163 KHTIIIGTAEWGGISA-----LEKLSVPKWE---KNSIVTIHYYNPFEF-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6.5%; Score 76.5; DB 3; Length 317; 21.1%; Pred. No. 1.5; tive 22; Mismatches 50; Indels 8
                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: Mathur, E., et al.
TITLE OF INVENTION: Carboxymethyl Cellulase from
TITLE OF INVENTION: Thermotoga Maritima
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    331400-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,889
                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: 08/518,615
FILING DATE: August 23, 1995
ATTORNEY/AGENT INFORMATION:
194 THMLFGSDRMELLAHLLGEKWMGP 217
                                      204 THQ -- GAEWVEGSEKWLGRKWGSP 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      204 THQ--GAEWVEGSEKWLGRKWGSP 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             194 THMLFGSDRMELLAHLLGEKWMGP 217
                                                                                                                                      Sequence 2, Application {\rm US/08951889} Patent No. 6008032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: FERRARO, GREGORY D. REGISTRATION NUMBER: 36,134 REFERENCE/DOCKET NUMBER: 331 TELECOMMUNICATION INFORMATION: 201-994-1700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       117 ASRE------LWMRVWSRNED-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 317 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 21.1
Matches 43; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA;
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE TYPE: PROTEIN
                                                                                                                                                                                                                                                                                                                                                                          NEW JERSEY
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINEAR
                                                                                                                                                                                                                                                                                                                                 STREET: 6 BECKEI
CITY: ROSELAND
STATE: NEW JERSI
COUNTRY: USA
                                                                                                               US-08-951-889-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-951-889-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                            Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
  ें
                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
73 LRHHLQI----PIHFPKDFL----SVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        247 IREMIQLPGARPILDPVDFLGLQDKIKEVPRPRKRLTELLLRTAT--EKPGPAEAARQAS 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14 VLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKG-LYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 WMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEA 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 497;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69; Indels
APPLICANT: MILLER, WALTER L.
APPLICANT: HARIKRISHNA, JENNIFER A.
APPLICANT: BLACK, STEPHEN M.
TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: MILLER, WALTER L.
APPLICANT: HARIRISHNA, JENNIFER A.
APPLICANT: BLACK, STEPHEN M.
TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZYMES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6.5%; Score 76.5;
25.4%; Pred. No. 3;
tive 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               UCAL-236/00US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/075,193
FILING DATE: 09-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NEELEY Ph.D., RICHARD L.
REGISTRATION NUMBER: 30,092
REFERENCE/DOCKET NUMBER: UCAL-236/06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494.7622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 4, Application US/08564090A Patent No. 5939318 Patent No. 5939318 5741703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: COOLEY GODWARD LLP
STREET: FIVE PALO ALTO SQUARE
                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-
```

ģ

```
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            73 LRHHLQI----PIHFPKDFL-----SVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            247 LREMIQLPGARPILDPVDFLGLQDKIKEVPRPRKRLTELLLRTAT--EKPGPAEAARQAS 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14 VLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKG-LYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122 WMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEA 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 497;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Application PC/TUS9406698
GENERAL INFORMATION:
APPLICANT: MILLER, WALTER L.
APPLICANT: HARIKRISHNA, JENNIFER A.
APPLICANT: BLACK, STEPHEN M.
TITLE OF INVENTION: CHOLESTEROL DISPOSAL FUSION ENZIMES NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSE: ROBBINS, BERLINER & CARSON STREET: 201 NORTH FIGUEROA STREET
                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/564,090A
FILING DATE: 02/05/96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 6.5%; Score 76.5; DB 2;
Best Local Similarity 25.4%; Pred. No. 3;
Matches 44; Conservative 29; Mismatches 69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: PCT/US94/06698 FILING DATE: FILED HEREWITH CLASSIFICATION: ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 90012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBB PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: RICHARD L. NEELEY, PH.I.
REGISTRATION NUMBER: 00.092
REFERENCE/DOCKET NUMBER: UCAI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5000
TELEFAN: 415-857-0663
                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
3000 EL CAMINO REAL
                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: 415-857-0663
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .... t .....rERISTICS:
TYPE: amino acids
TOPOLOGY: 11...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE: protein US-08-564-090A-4
                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                    PALO ALTO
CALIFORNIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOS ANGELES
CALIFORNIA
                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                  USA
                                                                                    94306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCT-US94-06698-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: L
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
```

```
73 LRHHLQI ----PIHFPKDFL-----SVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     247 LREMIQLPGARPILDPVDFLGLQDKIKEVPRPRKRLTELLLRTAT--EKPGPAEAARQAS 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14 VLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKG-LYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       305 ASRAWGL-RFFRSPQQVLPSPD----GRRAAGV--RLAVTRLEG-VDEATRA 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 WMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEA 174
                                                                                                                                                                                                                                                                                                                                                                                    DB 4; Length 497;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: MITTA, Masanori
APPLICANT: YAMAMOTO, Katsuhlko
APPLICANT: MORISHITA, Mio
APPLICANT: MORISHITA, Mio
APPLICANT: TSUNASAWA, Susumu
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/750,532
                                                                                                                                                                                                                                                                                                                                                                                                                                      29; Mismatches 69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street N.W., Suite 300
                                                                                                                                                                                                                                                                                                                                                                                 6.5%; Score 76.5; 25.4%; Pred. No. 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01095
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1994/130236
FILING DATE: 13-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1994/173912
FILING DATE: 26-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: BERLINER, ROBERT
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-224-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-977-1001
TELEFAX: 213-977-1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               United States of America
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-750-532-9
; Sequence 9, Application US/08750532
; Patent No. 5756339
                                                                                                                                                 INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 anino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                      44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: BROWDY, ROGER L. REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                             ; MOLECULE TYPE: protein PCT-US94-06698-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Washington
                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE: D.C.
                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
```

```
å
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                ó,
                                                                                                                                                                                                                                                                                                                                                                                                                     :: | :: | :: | :: | 1203 SSPEANATIVSVEMESEGGVKKTVTVEITINGTANETATIVVPVPKKAENIEVSGDHVIS 1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80 PIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPQSIL 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140 AAAEKAG--MSAE-QAQGLLEKIAT---------PKVKNQLKETTEAACR 177
                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             178 YG------AFGLPITVAHVDGQTHMLFGSDRMELLA-HLLGEKW 214
                                                                                                                                                                                                                                                                                                        DB 1; Length 1398;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Meyerwitz, Elliot M.
APPLICANT: Clark, Steven E.
APPLICANT: Clark, Steven E.
APPLICANT: Williams, Robert W.
TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
TITLE OF INVENTION: Transformed Plants, and Proteins
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
                                                                                                                                                                                                                                                                                                                                                69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: LEW PC COMPATIBLE FORMS
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COMPUTER: IBM PC COMPATIBLE
COURRENT SPECIFIC. PC-10S/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,553A
FILLING DATE: 06-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: S11va, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A 60886/RFT/RMS
TELECOMMUNICATION: NEORMATION:
                                                                                                                                                                                                                                                                                                                                                                                        50 DSGNKPPGLLPRKGLY--MANDLKLLRHHLQI-------
                                                                                                                                                                                                                                                                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                      6.4%; Score 75.5;
16.8%; Pred. No. 19
ative 39; Mismatche
REFERENCE/DOCKET NUMBER: MITTA=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 3, Application US/08473553A Patent No. 5859338 GENERAL INFORMATION:
                    TELECOMMUNICATION INFORMATION TELEPHONE: (202) 628-5197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEEX: 910 277299
INFORBATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
                                          TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 1398 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               523 amino acids
                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                              TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
NOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 39; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-473-553A-3
                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
14;
                                                                                                                                                                                                                                                                                                                                                                                 93 EKGSLSAMRFLTAVNL-----EHPEMLEKASRELWMRVWSRNEDITEPQSILAAAE-- 143
                                                                                                                                                                                                                                                                                                                                                                                                                                             144 KAGMSAEQAQGLL-----VKNOLKET 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       296 KLDVSDNHLTGLIPKDLCRGEKLEMLILSNNFFFGPIPFELGKCKSLTKIRIVKNLLNGT 355
                                                                                                                                                                                                                                                          36 NLOLRPSLITGIMKDSGNKPP---GLLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVML 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 TEAACRYGAFGLP-ITVAHV----DGQTHMLFGSDRMELLAHLLGEKWM-GPIPPAV 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  356 VPA----GLFNLPLVTIIELTDNFFSGELPVTMSGDVLDQI---YLSNNWFSGEIPPAI 407
                                                                                                                                                                                                            75;
                                                                                                                                                             Length 523;
                                                                                                                                                                                                            34; Mismatches 77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET: A EMBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400 STATE: California COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
                                                                                                                                                           6.2%; Score 73.5; DB 2; 21.8%; Pred. No. 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Comal, Lucio
APPLICANT: Dynlact, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Nooko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 32, Application US/08188582 Patent No. 5534410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 32:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 869 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tjian, Robert
                                                                                                                                                                                                            Conservative
                   STRANDEDNESS: unknowr

TOPOLGGY: unknown

MOLECULE TYPE: protein

US-08-473-553A-3
                                                                                                                                                           Ouery Match
Best Local Similarity
Matches 52; Conserva
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-188-582-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
```

Tue Apr

```
ДD
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------YRDPETLVFRDSSSWR--WA---DFTAHPRVLTVGDRTGVKMLD 412
                                                                                                                                                                                                                                                                                                                                                                 318 VWKFGKQWQPTLLQAMQVEKGATGISLSP----HLPGELAICSRSGAVCLWSPEDGLRQI 373
                                                                                                                                                                                                                                                                                                                                                                                                                                         92 LEKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQ 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         152 AQG-----LEKIATPKVKNQLKETTEAACRYGA-----FG-----LPITVAHVDG 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32 IMNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Wang, Edith
APPLICANT: Weinzlerl, Robert O.J.
TITLE OF INVENTION: TATA BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: UUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                            59;
                                                                                                                                                    DB 1; Length 869;
                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT STREET: 4 Embarcadero Center, Suite 3400 CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: PATEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
ADDIT APPLICATION: ABDIT APPL
                                                                                                                                                                                                                    82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A-57650-2/AJT/RAO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                462 QFSLYLVDERLPLVPML---KWNHGLPSPLLLARL 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      193 QTHMLFGSDRMELLAHLLGEKWMGPIP-PAVNARL 226
                                                                                                                                            6.2%; Score 73.5; D 20.5%; Pred. No. 15; Live 30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: OSMMAN, RICHARD A
REGISTRATION NUMBER: 36,627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 32, Application US/08646715
Patent No. 5637686
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tjian, Robert
Comai, Lucio
Dynlact, Barlan D.
Hoey, Timothy
Ruppert, Slegfried
Tanese, Naoko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIF: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE/DOCKET NUMBER: A-TELECOMMUNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEPHONE: (415) 781-1989
TELEFAX: (415) 388-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 32:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 869 amino acids
                                                                                                                                                                              Best Local Similarity 20.58 Matches 44; Conservative
           protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear
; MOLECULE TYPE:
US-08-188-582-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-646-715-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 374 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
14;
                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                         374 -----YRDPETLVFRDSSSWR--WA---DFTAHPRVLTVGDRTGVKMLD 412
                                                                                                                                                                                                                                                                                                                                                                                         152 AQG-----LLEKIATPKVKNQLKETTEAACRYGA-----FG----LPITVAHVDG 192
                                                                                                                                                                                                                                                                                                                                                                                                                          413 TQGPPGCGLLLFRLG------AEASCQKGERVLLTQYLGHSSPKCLPPTLHLVCT 461
                                                                                                                                                                                                                                        318 VWKFGKQWQPTLLQAMQVEKGATGISLSP----HLPGELAICSRSGAVCLWSPEDGLRQI 373
                                                                                                                                                                                                                                                                                             92 LEKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQ 151
                                                                                                                                            59; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                            32 IWNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 980;
                                                                                                DB 1; Length 869;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                              82; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Meyerowitz, Elliot M.
APPLICANT: Clark, Steven E.
APPLICANT: Williams, Robert W.
TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
TITLE OF INVENTION: Transformed Plants, and Proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
PAPLICATION NUMBER: US/08/473,553A FILING DATE: 06-JUN-1995 CLASSIFICATION: 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.2%; Score 73.5; DB 2; 1
21.8%; Pred. No. 18;
tive 34; Mismatches 77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      193 QTHMLFGSDRMELLAHLLGEKWMGPIP-PAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A-60886/RFT/RMS
                                                                                           Query Match 6.2%; Score 73.5; D. Best Local Similarity 20.5%; Pred. No. 15; Matches 44; Conservative 30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPOTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-473-553A-6; Sequence 6, Application US/08473553A; Patent No. 5859338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 06-JUN 1995
CLASSIFICATION: 800
ATTORNEY FAGENT INFORMATION:
NAME: S11va, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-601
TELECHONE: (415) 781-1989
TELECHONE: (415) 38-3249
TELECHONE: (15) 38-3249
TELECK: 910 277299
TELECK: 910 277299
INFORMATION FOR SEQ ID NO: 6: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 980 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: United States ZIP: 94111-4187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 21.8*
Matches 52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein US-08-473-553A-6
; MOLECULE TYPE: protein US-08-646-715-32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Ho
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              San Francisco
California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
```

Oy Ob	36 NLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVML 92	
[*] Oy Ob	93 EKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPQSILAAAE 143 :	
Qy	144 KAGMSAEQAQGLLEKIATPKVKNOLKET 171 .	
λ̈́δ	TEAACRYGAFGLP-ITVAHVDGQTHMLFGSDRMELLAHLLGEKWM-GPIPPAV 222	
QQ		
A S S S S S S S S S S S S S S S S S S S	BESULT 15 US-06-173-52A-2 US-06-173-52A-2 US-06-173-52A-2 PAPALICANT WARD GENERAL INFORMATION: APPLICANT: Clark, Steven E. APPLICANT: Williams, Pobert W. ITILE OF INVENTION: Plant Clavatal Nucleic Acids, ADRESSEE Flehr, Hobbach, Test, Albritton & Herbert STREET Four Embarcadero Center, Suite 3400 CITY: San Frantisco STATE: California COUNTR: United States INFORMATION FOR ENDER PRANTION: MEDIUM TYPE: FLOPPY disk COMPUTER RADABLE FORM: MEDIUM TYPE: FLOPPY DISK COMPUTER: INH FC compatible COMPUTER: INH FC compatible COMPUTER: INH FC compatible COMPUTER: SALVA, ROSH APPLICATION NUMBER: 38,304 RECESTRATION NUMBER: 38,304 RECESTRATION NUMBER: 38,304 RECESTRATION NUMBER: 38,304 REEFERAX: (415) 399-3249 TELEFRAX: (415) 399-3249 TELEFRAX: 910 277299 TELEFRAX: 910 277299 TELEFRAX: MINORARISICS: LENGTH FORMATION COMPATION: FELENCALE TYPE: Anno acid TYPELCAL TYPE: PLOCEIN WOLKER TYPE: PLOCEIN	
On Be Ma	Ouery Match 6.2%; Score 73.5; DB 2; Length 985; Best Local Similarity 21.8%; Pred. No. 18; Matches 52; Conservative 34; Mismatches 77; Indels 75; Gaps 14;	
QQ Dp	36 NLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVML 92	
Oy Dp	93 EKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPOSILAAAE 143 	

144 KAGMSAEQAQGLLVKNQLKET 171	370 KLDVSDNHLTGLIPKDLCRGEKLEMLILSNNFFFGPIPEELGKCKSLTKIRIVKNLLNGT 429	172 TEAACRYGAFGLP-ITVAHVDGQTHMLFGSDRMELLAHLLGEKWM-GPIPPAV 222	430 VPAGLENLPLVTIIELTDNFFSGELPVTMSGDVLDQIYLSNNWFSGEIPPAI 481
144	370	172	430
Qy	qa	Οy	qq

Search completed: April 3, 2001, 11:38:40 Job time: 23 sec

		1		

	,			
	,			
			<i>5</i> .	
				,
			¥)	
		2	,	
				•
•	•			

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

April 3, 2001, 11:38:20 ; Search time 15.64 Seconds (without alignments) 981.173 Million cell updates/sec

Title: Perfect score:

US-09-441-723-1 1185 1 MGPLPRTVELFYDVLSPYSW......AHLLGEKWMGPIPPAVNARL 226 Sednence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

195891 seqs, 67900655 residues Searched:

195891 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_66:* Database :

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	hypothetical prote	hypothetical prote		2-hydroxychromene-	2-nitrotoluene dio	glutathione transf	2-hydroxychromene-	ABC transporter, A	flavoprotein A hom	pitrilysin (EC 3.4	probable isomerase	threonine dehydrat	iO.	probable glycosyl	probable tRNA delt		catalase (EC 1.11.	kinesin-related pr	hypothetical prote	secretogranin II -	DNA-directed DNA p	hypothetical prote	protein kinase CDC	hypothetical prote	topoisomerase I -	catalase (EC 1.11.	heat-stable entero	rRNA methylase Spo	conserved hypothet
		T27747	T34201	G83629	S72164	T31286	S17164	C55552	H72425	C69127	SNECPI	I49343	T51712	A35296	E83375	B71301	S54595	A40367	S09748	T51169	S02180	DJAD12	875896	A48144	E69107	A70358	E69647	охнинх	C75301	F83215
	DB	1	~	~	~	~	~	~	7	~	Н	7	~	~	~	7	7	7	П	7	7	-	~	~	7	7	7	٦	7	7
	Query Match Length DB	226	225	195	194	197	33	203	581	410	962	199	592	613	926	316	971	488	700	263	619	1061	274	705	402	540	547	1073	230	333
æ	Query Match	27.7	25.2	16.7	13.9		10.5	8.0		7.6	•	7.5	•	7.5	•	7.1	•	•	7.0	7.0	6.9	6.9	6.8	6.8	6.8	6.8	6.7	6.7	٠.	9.9
	Score	328.5	298.5	198	164.5	140	125	95	91	90.5	89.5	68	88	88.5	87			83.5	83.5	82.5	81.5		81	80.5	80	80		79.5	79	78.5
	Result No.		2	e	4	S	9	7	8	6	10	11	12	13	14	15	16	17	. 18	19	20	. 21	22	23	24	. 25	26	27	28	29

alanine dehydroger	cellulase (EC 3.2.	branched-chain ami	hypothetical prote	probable o-stalogl	cation-transportin	botulinum toxin no	transcription regu	probable transposa	hypothetical prote	hypothetical prote	transport ATP bind	hypothetical prote	ferredoxinNADP+	ferredoxinNADP+	Appropriate overlage
G71648	T06060	G69352	C72784	A71545	C69069	I40817	T17401	T44628	C64454	G75008	D72083	T40215	A40487	JT0751	UB2362
~	~	~	7	~	7	7	7	~	~	7	7	7	-	н	c
383	493	248	313	338	910	1162	1479	482	634	253	629	184	497	498	6 7 0
9.9	9.9	9.9	9.9	9.9	9.9	9.9	9.9	6.5	6.5	6.5	6.5	6.5	6.5	9	U
	78.5 6.6	78 6.6	78 6.6	78 6.6	78 6.6	78 6.6	78 6.6		77.5 6.5		77 6.5			76.5 6.5	

ALIGNMENTS

N RESULT T34201

hypothetical protein D2024.7 - Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Date: 29-oot-1999 #sequence_revision 29-Oot-1999 #text_change 29-Oct-1999 C;Accession: T34201 R;Du, Z.; Gattung, S: Submitted to the EMBL Data Library, November 1995 A;Description: The sequence of C. elegans cosmid D2024. A;Reference number: 221488

```
Best Local Sim
Matches 54;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Gene: nahD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C. Species: Pseudomonas aeruginosa
C. Species: Pseudomonas aeruginosa
C. Species: Pseudomonas aeruginosa
C. Date: 15-Sep-2000
R. Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N. V.
Natura 406, 959-964, 2000
A. Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathon A; Reference number: A82950
A; Tatus: preliminary
A; Mossiques: 1-195 < STO>
A; Residues: 1-195 < STO>
A; Residues: 1-195 < STO>
A; Cross-references: GB:AE004450; GB:AE004091; NID:g9945943; PIDN:AAG03508.1; GSPDB:GN001
A; Scenetics:
A; Genetics:
A; Genetics:
                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ς.
Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             G83629
hypothetical protein PA0118 [imported] - Pseudomonas aeruginosa (strain PA01)
                                                                                                                                                                                                                                                                                                                                                                                           120 GLWHRFYAYGKPIFTKSQ---VAEVLRDLHVKDVDELVWMSDSAEVKNIIJRENTDEAIGN 176
                                                                                                                                                                                                                                                                                                                                                                  PRKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASR 119
                                                                                                                                                                                                                                                                                                                                                                                                                                               ELWMRVWSRNEDI-TEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRY 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LYMANDLKLLRHHLQIPI----HFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASR 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 RYTLHDLARYAKRYGVPLAFNPAFPINTLTLM-----RGAQGYLGG-----EGFQPYLK 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120 ELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 PLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL---L 59
                                                                                                                                                                                                                                                                                                          4 LPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKG 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MSKQIEFFFDFGSPTTYLAWTQLPRIAAAHGASIAWRPMLLGGVFKATGNHSPIEVPAKG 60
                                                                                                                                                                                                                                                 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 195;
A;Accession: T34201
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Rosidues: 1-225 < DUZ>
A;Conetics: EMBL:U41011; PIDN:AAA82289.1; CESP:D2024.7
A;Genetics:
A;Genetics: A;Actions: 51/1: 177/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Score 198; DB 2; Length 19; Pred. No. 7.2e-10; 41; Mismatches 93; Indels
                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             179 GAFGLP---ITVAHVDGQT-HMLFGSDRMELLAHLLGEKWMGPI 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25.2%; Score 298.5; DB 2
33.0%; Pred. No. 2.1e-18;
iive 43; Mismatches 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180 AFGLPITVAHVDGQTHMLFGSDRMELLAHLL 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16.7%;
26.1%;
                                                                                                                                                                                             Query Match
Best Local Similarity 33.09
Marches 74; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 26.1%
Matches 55; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                        09
                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                               ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
Cydecession: S72164
RyBrito, B.; Palacios, J.M.; Ruiz-Argueeso, T.; Imperial, J.
Biochim. Biophys. Acta 1308, 7-11, 1996
A; Hitle: Identification of a gene for a chemoreceptor of the methyl-accepting typ A; Hitle: Identification of a gene for a chemoreceptor of the methyl-accepting typ A; Reference number: S72162, MUID:96328256
A; Accession: S72164
A; Molecule type: DM
A; Mossium: S72164
A; Residues: 1-194 < CREI>
A; Cross-references: EMBL:U23040; NID:9780654; PIDN:AAC44312.1; PID:9780657
A; Cross-references: EMBL:U23040; NID:9780654; PIDN:AAC44312.1; PID:9780657
C; Genetics:
A; Genome: plasmid
C; Superfamily: 2-hydroxychromene-2-carboxylate isomerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2-iitrotoluene dioxygenase (EC 1.14...) Rieske iron-sulfur component - Sphingomo C; Species: Sphingomonas aromaticivorans
C; Species: Sphingomonas aromaticivorans
C; Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C; Accession: T31286
R; Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, submitted to the EMBL Data Library, July 1998
A; Reference number: 220992
A; Reference number: 220992
A; Reference number: 220992
A; Accession: T31286
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-197 < ROM>A; Cross references: EMBL: AF079317; NID: 93378261; PID: 93378427; PIDN: AAD04010.1
          pv.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7;
2-hydroxychromene-2-carboxylate isomerase homolog - Rhizobium leguminosarum C;Species: Rhizobium leguminosarum bv. viciae
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 29-Sep-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106 --EAIFNAIWSAPAPLATAAEVAAVLGAAGLDAEE---LAERMDEPAAQDLLDEATANAV 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65 YMANDLKLLRHHLQIPIHFPKDFLSV-----MLEKGSLSAMRFLTAVNLEHPEMLEK 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       117 ASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAAC 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 RKGLYMANDLKLLRHHLQIPIHFPKDF-----LSVMLEKGSLSAMRFLTAVNLEHPEML 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 RTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL-LPRKGL 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 LPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPG---LLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 194;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Genome: plasmid pNL1
C;Superfamily: 2-hydroxychromene-2-carboxylate isomerase
C;Keywords: oxidoreductase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11.8%; Score 140; DB 2; 23.1%; Pred. No. 6.9e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13.9%; Score 164.5; DB 25.2%; Pred. No. 5.4e-07; Live 34; Mismatches 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         177 RYGAFGLPITVAHVDGQTHMLFGSDRMELL-AHL 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161 SRGVFGAPTLFV----GDEMFFGNDRLHFMQGHL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23.1%; Pred. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity 25.2'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity
Matches 50; Conserv
```

DD 61 -KIKVMKADLERWAER	: : : : : : : : : : : : : : : : : :	Qy 123 MRVWSRNEDITEPQSILAAAE
Qy 115 EKASRELWMRVWSRNE : 1:1: Db 110DAYRIWGGGI	EKASRELWMRVWSRNEDITEPQSILAAABKAGMSAEQAGGLLEKIATPKVKNQLKETTEA 174 	181
Qy 175 ACRYGAFGLPITVAHV : : : Db 162 AIQRGVYGAPLMFV	AHVDGQTHMLFGSDRWELLAHLL 210	SULT
RESULT 6 \$17164 glutathione transferase (E C:Species: Rattus norvegic C;Date: 21-Nov-1993 #seque	RESULT 6 \$17164 \$17164 State transferase (EC 2.5.1.18) 13 - rat C:Species: Rattus norvegicus (Norway rat) C;Date: 21-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997	ABC transporter, ATP-binding prot C;Species: Thermotoga maritima C;Date: 11-Jun-1999 #sequence_rev C;Accession: H72425 R;Nelson, K.E.; Clayton, R.A.; G; Garrett, M.M.; Stewart, A.M.; Col
C; Accession: S17164 R; Harris, J.M.; Meyer, D.J.; Coles, B.; Kett. Biochem. J. 278, 137-141, 1991 A; Title: A novel glutathione transferase (13 A; Reference number: S17164; MUID:91354194 A; Recession: S17164 A; Status: preliminary A; Molecule type: protein	J.; Coles, B.; Ketterer, B. 1991 one transferase (13-13) isolated from the matrix of rat liver m 4; MUID:91354194	C.M. Nature 399, 323-329, 1999 A)Title: Evidence for lateral gen A,Reference number: A72200; MUID A,Accession: H72425 A,Status: preliminary A,Molecule type: DNA A,Residues: 1-581 <arn></arn>
A; Residues: 1-33 <har> C; Keywords: transferase Query Match</har>	10.5%; Score 125; DB 2; Length 33;	A.Cross-references: GB:AE001691; A.Experimental source: strain MSI C.Genetics: A.Gene: TW0043 C.Superfamily: Escherichia coli
Matches 23; Conservative Oy 3 PLPRTVELFYDVLSPYSWL	4; Mismat GFEILCRYONIWN	Query Match 7.7%; Best Local Similarity 19.0%; Matches 35; Conservative
Db 2 PAPRVLELFYDVLSPY RESULT 7	SPYSXLGFEVLXRYQHLXN 33	Oy 28 RYONIWNINLO :: : Db 343 RFENVWFSYDGKNWVLKDINLD
C55552 2-hydroxychromene-2-carbox C;Species: Pseudomonas put C;Date: 08-Sep-1995 #seque	boxylate isomerase - Pseudomonas putida plasmid NAH7 putida putida 1995 #text_change 29-Sep-1999	Qy 69 DLKLLRHHLQIPIHFPKD :: :: : Db 403 EIPLLEYNLKLVRKQIAAVPQD
 C; Accession: C55552 R; Eaton, R.W. J. Bacterill 176, 7757-7762, 1994	catabolic nathw	Qy 116 KASRELWMRVWSRNEDITEPQS : : : : : : : : : : : : : : : : : : :
A) Accession (5555) A) Accession (5555) A) Accession (5555) A) Accession (5555) A) Status; preliminary		170
A:Residues: 1-203 < EAVY A:Residues: 1-203 < EAVY A:Cross references: GB:U00 C:Genetics: A:Gene: nabb A:Genome: plasmid	U09057; NID:9483790; PIDN:AAA66358.1; PID:9483793	RESULT 9 C69127 flavoprotein A homolog (II) - MeC; Species: Methanobacterium ther C; Species: Methanobacterium ther
C;Superfamily: 2-hydroxyc	2-carboxylate isomerase	C; Accession: C69127 R; Smith, D.R.; Doucette-Stamm, L.; Qiu, D.; Spadafora, R.; Vicair
Query Match Best Local Similarity 24.2%; Matches 51; Conservative 3	tch 8.0%; Score 95; DB 2; Length 203; sal Similarity 24.2%; Pred. No. 0.53; 51; Conservative 35; Mismatches 97; Indels 28; Gaps 8; versynvisovswigfeticRyonywiniolRPSLITGIMKDSGNKPPGLLPRKGLY 65	K1, S.; CNUTCh, G.M.; Danlels, C. J. Bacteriol. 179, 7135-7155, 19 A;Title: Complete genome sequenc A;Reference number: A69000; MUID A;Accession: C69127
3 99	DLKVKLDY EKASRELW	A;Status: preliminary; nucleic a A;Molecule type: DNA A;Residues: 1-410 cMTH3 A;Cross-references: GB:AE000809; A;Experimental source: strain De
Db 63 LKVDLQRWAQLYGIP	PLVFPANYNSRRMNIGFYYSGAEAQAAAYVNVVF 110	C;tenetics: A;Gene: MTH220

```
Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; otton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardso
                                                                                                                                                                                                                                                                                                                                                    Jene transfer between Archaea and Bacteria from geno
:D:99287316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; GB:AE000512; NID:94980517; PIDN:AAD35137.1; PID:9
SB8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABC transporter mdlA; ATP-binding cassette homolog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wethanobacterium thermoautotrophicum (strain Delta Parmoautotrophicum
revision 05-Dec-1997 #text_change 24-Sep-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            L.A.; Deloughery, C.; Lee, H.; Dubols, J.; Aldredge
ire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwe
C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nce of Methanobacterium thermoautotrophicum Delta H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; GB:AE000666; NID:92621265; PIDN:AAB84726.1; PID:
Delta H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                    evision 11-Jun-1999 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DFL-----SVMLEKGSLSAMRFLTAVNLEHP-EMLE 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SILAAAEKAGM----SAEQAQGLLEKIATPKVKNQLK 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QLRPSLITGIMKDSGNKPPGL-----LPRKGLYMAN 68
                                                                                                                                                                                                              otein - Thermotoga maritima (strain MSBB)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Score 91; DB 2; Length 581;
; Pred. No. 4.5;
44; Mismatches 67; Indels 38;
                                                                              RMELLAHLLG 211
                                                                                                      |::| :|
RLFMLESAMG 192
```

3

```
position: 61
                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pltfilysin (EC 3.4.24.55) precursor [validated] - Escherichia coli
Nictorate names: anodopetidase Pi; proteinase III
Cispecia: Bacherichia coli
Airita: The complete genome sequence of Escherichia coli K-12.
Airita: The complete genome sequence of Escherichia coli K-12.
Airita: The complete genome sequence of Escherichia coli K-12.
Airita: The complete genome sequence of Escherichia coli K-12.
Airita: Mulei Dia Cispecia: Bacherichia coli K-12.
Airita: Anodole I complete genome sequence of Escherichia coli K-12.
Airita: Dia Cispecia: Bacherichia coli Relation not shown
Airita: Dia Cispecia: Bacherichia coli Bacherichia coli Cispecia: Bacherichia Airita: Bacherichia coli Escherichia Airita: Dia Cispecia: Bacherichia Cispecia: Bacherichia Cispecia: Bacherichia Cispecia: Bacherichia Cispecia: Bacherichia Cispecia: Bacherichia Cispecia: Bacherichia: Airita: Analysis of the requilatory region of the protease III (ptr) gene of Escherichia
Airita: Analysis of the requilatory region of the mature protein, was confir Airita: Complete muclectide sequence of the Escherichia coli ptr gene encoding protease
Airita: Complete muclectide sequence of the Escherichia coli ptr gene encoding protease Airita: Airichia: Airita: Ai
                                                                                                                                                                                                                       16;
                                                                                                                                                                                                                                                                                                                                                                                      ----PPG----LLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRF----L 103
                                                                                                                                                                                                                                                                                                                                                                                                                                       153 MLHWPDSMFTLLEEEGILFSND--AFGQHLCISKRFDKDVPEAVLMD---AAMKFYANLL 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104 TAVNLEHPEMLEKAS--RELWM------RVWSRNEDITEPQSILAAAEKAGMSA 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99 EVLKRY---GSEIICTAKAAEGLROHYSIPQDTPMQ---TVKTGDSTDLGGKTLTFLEAP 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EQAQGLLEKIATPKVKNQLKETTEAACRYGAFGLPIŢVAHVDGQTHMLFGSDRMELLAHL 209
                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPLS---PLVLRKFSEVKELGLLEKIGMIAPSHGQIW-----TEPLKIIAA----YT
                                                                                                                                                                                                                       81;
                                                                                                                                                         DB 2; Length 410;
                                                                                                                                                                                                                                                                     DVLSPYSWLGFEILC------RYQNIWNINLQLRPSLITGIMKDSGNK-
                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                             82;
                                                                                                                                                                                  ; Pred. No. 3.2; 31; Mismatches
                                                                                                                                                      7.6%; Score 90.5; 25.1%; Pred. No. 3.2
                                    Superfamily: Methanobacterium flavoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              210 LGEK------ WMGPIP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   310 LESKAVFIGSPTMFNGPFP 328
                                                                                                                       Ouery Match
Best Local Similarity 25.18
                                                          C: Keywords: flavoprotein
A>Start codon: TTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    253
                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

```
A Pathway: protein degradation
C; Superfamily: insulysin
C; Superfamily: insulysin
C; Superfamily: insulysin
C; Superfamily: periplasmic space; protein degrees, respondent of Signal sequence #status predicted (SIG>
F; 1-23/Domain: signal sequence #status experimental <AMT>
F; 88,92/Finding site: zinc (His) #status experimental
F; 91/Active site: Glu #status experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Species: Pseudomonas sp. (Strain Clo)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 18-Sep-1998
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 18-Sep-1998
C;Accession: 14934; S27640
G;Benome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.
J. Bacteriol. 175, 6890-6901, 1993
A;Title: Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: characterior umber: A49343; MUID:94042852
A;Accession: 149343
A;Accession: 149343
A;Accession: 149343
A;Accession: 1499 cbb>
A;Cross-references: GB:M60405
C;Genetics:
C;Function:
A;Description: endopeptidase degrades small peptides [validated; MUID:92237263]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      threonine dehydratase/deaminase [imported] - Arabidopsis thallana C;Species: Arabidopsis thaliana (mouse-ear cress) C;Species: Arabidopsis thaliana (mouse-ear cress) C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000 C;Accession: T51712 R;Mourad, G.S.; Emerick, R.W.; Marion, A.L.; Smith, A.M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                731 WCRNKDVVVDKKQSVIFEKAGNSTDSALAAVFVPTGYDEYTSSAYSSLLGQIVQPWFYNQ 790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 MANDLKLLRHHLQIPIHFPKDFLSVMLEKGSL--SAMRFLTA-VNLEHPEMLEKASRELW 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 MRVWSRNEDITEPQSILAA--AEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111 NAVW--GDGIAPDLESLPALVSEKLGWDRSAFE---DFISSDAATERYDEQTHAAIERKV 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                       23; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8 VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    168 LKETTEAACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPA 221
                                                                                                                                                                                                                                                                                                                                                                                  Length 962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 7.5%; Score 89; DB 2; Length 199; Best Local Similarity 24.2%; Pred. No. 1.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Gene: doxJ
C;Superfamily: 2-hydroxychromene-2-carboxylate isomerase
                                                                                                                                                                                                                                                                                                                                                                               ch 7.6%; Score 89.5; DB 1; 1
1 Similarity 26.3%; Pred. No. 12;
30; Conservative 13; Mismatches 48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   probable isomerase doxJ - Pseudomonas sp. (strain C18)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           166 FGVP-TMFLGD---EMWWGNDRLFMLENAVG 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 FGLPITVAHVDGQTHMLFGSDRMELLAHLLG 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51;
```

÷

Tue Apr

1

```
Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73 LRHHLQIPIHFPKDFLSVMLEKGSLSAM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best_Local Similarity 23.39
Matches 41; Conservative
         R;Stover, C.K.; Pham, X.Q.;
adman, S.; Yuan, Y.; Brody,
..., Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
                                                                                                                                                                                                                     -926 <STO>
                                                                                                                                                                                                  A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                      A; Gene: PA2162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Species: Bos prinigenius taurus (cattle)
C; Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 24-Sep-1999
C; Accession: A35296
C; Accession: A35296
R; Fischer-Colbrie, R.; Gutierrez, J.; Hsu, C.M.; Iacangelo, A.; Eiden, L.E.
J. Biol. Chem. 265, 9208-9213, 1990
A; Title: Sequence analysis, tissue distribution and regulation by cell depolarization, A; Reference number: A35296; MUID: 90264409
A; Accession: A35296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    probable glycosyl hydrolase PA2162 [imported] - Pseudomonas aeruginosa (strain PAO1) (;Species: Pseudomonas aeruginosa C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000 C;Accession: B83375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             φ
..
                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          105 -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLQPVFSFKLRGAYNMMVKLPADQLAK 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  378 EQ------ELEIPVE-PED------ISEVDLDHPDLFQN---- 403
                                                                                                                                                                                                                                                                                                                                                              55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 52 PPKLPLPR------LKVSPNSLQ----YPAGYLGAVPERTNEAENGSIAEAMEYLTNI 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RIVELFYDVLSPYSWLGFEILCRYQNIWNI-----NLQLRPSLITGIMKDSGNKPPGLLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A Status: preliminary
A Molecule type: mRNA
A: Residues: 1-613 <PISS
A: Residues: 1-613 <PISS
A: Cross-references: GB:J05468; NID:g163709; PIDN:AAA30760.1; PID:g163710
C; Superfamily: secretogranin II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2; Length 613;
                                                                                                                                                                                                                                                                       Length 592;
                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLL--EKIATPK 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       160 GVICSSAGNHAQGVALSASKLGCTAVIVMPVTTPEIKWQAVENL--
                                                                                                                                                                                                                                                                                                                    57;
Submitted to the EMBL Data Library, October 1998
A;Reference number: 225433
A;Recession: T51712
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-592 <MOU>
A;Cross-references: EMBL:AF096281; PIDN:AAC97936.1
C;Genetics:
                                                                                                                                                                                                                                                                       DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                  Query Match 7.5%; Score 89; DB 2
Best Local Similarity 22.8%; Pred. No. 6.9;
Matches 49; Conservative 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.5%; Score 88.5; DB ilarity 23.6%; Pred. No. 8; Conservative 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      secretogranin II precursor - bovine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142 AEKAGMSAEQAQG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 39; Conserva
                                                                                                                                                                                               A; Gene: OMR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 404
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunisti
A;Reference number: A82950
A;Accession: E83375
A;Status: preliminary
                                                                                                                                                                                        A;Residues: 1-926 <STO>
A;Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05550.1; GSPD
A;Experimental source: strain PAO1
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     probable tRNA delta(2)-isopentenylpyrophosphate transferase (miaA) - syphills spi
C;Species: Treponema pallidum subsp. pallidum (syphills spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: B71301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Residues: 1-316 <CCL>
A;Cross-references: GB.AE001238; GB.AE000520; NID:g3322928; PIDN:AAC65611.1; PID: A;Experimental source: strain Nichols
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete. A;Reference number: A71250; MUID:98332770
A;Recession: B71301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RiFraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, rson, J.; Khalak, H.; Rtchardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Status: preliminary; nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKDFLSVMLEKGSLSAMRFLTAVN-LEHPEMLEKASRELWMRVWSRNE----DITEPQSI 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 LAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGL----PI--TVAHVDG 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          412 VLAGSLAGDLENLAGGLL-RVARADLAS--RDLTLGGIRRALFOLLARFPVYRTYAGACG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80 DVCDPYE--EYNVFRFQQAVYGI----VPSIL-----RAHKVPIIVGGTGLYLD---AV 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----- RFLTAVNL- 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           125 LRQYALVPVE-RNQALRASLRGASLSHMRAVYFSLKDSHAVHNKTDLEDPARLMRAIEIA 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13 DVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   109 ----EHPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIA 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            193 QT-----PIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.1%; Score 84.5; DB 2; Length 316; 23.3%; Pred. No. 7.4; tive 30; Mismatches 60; Indels 4
                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 7.3%; Score 87; DB 2; Length 926; Best Local Similarity 26.4%; Pred. No. 18; Matches 46; Conservative 27; Mismatches 59; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Gene: TP0637
C;Superfamily: delta(2)-isopentenylpyrophosphate transferase
```

Search completed: April 3, 2001, 11:39:25 Job time: 65 sec

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

April 3, 2001, 11:38:45 ; Search time 10.33 Seconds (without alignments) 706.531 Million cell updates/sec

US-09-441-723-1 1185 1 MGPLPRTVELEYDVLSPYSW......AHLLGEKWMGPIPPAVNARL 226

Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

88757 seqs, 32294092 residues Searched:

88757 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SwissProt_39:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		æ			SUMMAKIES	0	
Result No.	Score	Query Match	Length	DB	ΩI	Description	uc
	869	73.3	225	; -	GTK1 RAT	14a 874473	rattue norv
2 1	328.5	27.7	226	- 1	VS21 CARET.		
m	298.5	25.2	225	-	YYS7 CAEEL	-	caenorhabdi
4	95	8.0	203	н	NAHD_PSEPU	051948 pset	pseudomonas
S	89.5		962	Н	PTRA_ECOLI		escherichia
9	89	7.5	212	П	NAHD_PSESP		pseudomonas
7	88.5		613	٦	SG2_BOVIN		bos taurus
æ	84.5	7.1	316	٦	MIAA_TREPA		treponema p
σ		7.1	971	٦	YM8P_YEAST		saccharomyc
10	83.5	•	488		CATA_LISSE		listeria se
11	83.5		700		NCD_DROME		drosophila
12			619	П	SG2_RAT		rattus norv
13	81.5	6.9	1061	7	DPOL_ADE12	_	human adeno
14	81		274	П	YG73_SYNY3		synechocyst
15	80.5	•	339	П	PURK_BRUME		brucella me
16	80.5		705	П	CC5_YEAST	P32562 sac	saccharomyc
17	•		1073	7	HSER_PIG		sus scrofa
18	80		540	П	TOP1_AQUAE		aquifex aeo
19	æ		1162	П	BXEN_CLOBO		clostridium
20	•	•	547	П	CATX_BACSU		bacillus su
21	79.5	•	1073	Н	HSER_HUMAN	P25092 homo	o sapien
22	•	•	383	٦	PNAA_RICPR		rickettsia
23	_		1162	Н	BXEN_CLOBU	_	clostridium
24	77.5	•	634	П	YC36_METJA	_	methanococc
25	7		484	Н	ADRO_RAT		rattus norv
56	76.5		962	П	YBX7_SCHPO		schizosacch
27			1056	Н	DPOL_ADE02	_	numan adeno
28	٠		1193	Н	DPOL_ADE04	_	human adeno
. 59	97	6.4	406	7	PDK3_HUMAN	_	homo sapien
30	97	6.4	646	П	NODQ_RHIS3		r nodq bifu
31	97	6.4	2208	۲	POLN_MANCV	Q69014 mand	manchester
32	75.5	6.4	744	٦	GYRA_AQUAE		aquifex aeo
33	75.5	6.4	8	1	GLND_KLEPN	P41393 kle	klebsiella

							aquifex aeo				
P04495	P57030	P09127	040545	P89518	P47788	P40020	067718	003477	P25857	P00544	064948
POL_ADE05	COLD_NEIMA	HEMX_ECOLI	KPYA_TOBAC	OE66_NPVLS	MEPD_PIG	PIP1_YEAST	SECA_AQUAE	LAFT_VIBPA	G3PB_ARATH	FGR_FSVGR	LON1_ARATH
Ω	_			_	_			_	_	_	_
u u	-	-	Н	٠.				٠.			
Н	Н	393 1	593 1	631	989	925 1	984	285	402	545	888
1056 1	231 1						6.3 984]				
1056 1	6.3 231 1	6.3	6.3	6.3	6.3	6.3		6.3	6.3	6.3	6.3

ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.
                                                                                                                                    PRKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASR 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121
                                                                                                                                                                                  WMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 GPAPRVLELFYDVLSPYSWLGFEVLCRYQHLWNIKLKLRPALLAGIMKDSGNQPPAMVPH 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
                                                    GPLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPR 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MGPLPRTVELFYDVLSPYSWLGFEILCRYQNIW-NINLQLRPSLITGIMKDSGNKPPGLL 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7;
                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBMITTER (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 226;
                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                          15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
HYPOTHETICAL 26.5 KDA PROTEIN ZK1320.1 IN CHROMOSOME II.
                                                                                                                                                                                                                                                                   182 GLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DBA39A5994300164 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AFGLP-ITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIP 219
   ; Pred. No. 1.5e-68; 38; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27.7%; Score 328.5; DB 1; 33.0%; Pred. No. 1.3e-21; iive 47; Mismatches 94;
                                                                                                                                                                                                                                                                                                                                                                 226 AA
                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, 246934; CAA87039.1; -. WORMPEP; ZK1320.1; CE01698. Wortherical protein. SEQUENCE 226 AA; 26482 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 33.09
Matches 73; Conservative
Best Local Similarity 69.3
Matches 156; Conservative
                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                               YS21_CAEEL
Q09652;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Berks M.;
                                                                                                                                                                                                                                                                                                                                                YS21_CAEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                qq
                                                                                                                                                                                                                pp
                                                                                                                                                                                                                                                                                                                                                                 Dp
                                                                                                                   à
                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7;
                                                                                                                                                                                                                                          Du Z., Gattung S.;
Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
-! SIMILARITY: BELONGS TO THE GST SUPERFAMILY. KAPPA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASR 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIKEKYMHKDLLFSAQYWGIPFRLPKDYTNMMLNTSSIVPQRILVASQLRDNVLMEDVAR 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120 ELWMRVWSRNEDI-TEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRY 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL---L 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 PNRKVVKFFFDVISPYSYFGFEGITRHRSVWKTPIQMKPFFFAGVVRHTEN--PGLPLRI 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
2-HYDROXYCHROMENE-2-CARBOXYLATE ISOMERASE (HCCA ISOMERASE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                               15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
HYPOTHETICAL 26.0 KDA PROTEIN D2024.7 IN CHROMOSOME IV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23D0A3D6762B7232 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179 GAFGLP---ITVAHVDGQT-HMLFGSDRMELLAHLLGEKWMGPI 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Score 298.5; DB 1;
; Pred. No. 5.2e-19;
43; Mismatches 92;
                                225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             203
                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LEJUENCE FROM N.A.
STRAIN-G7 / ATCC 17485;
MEDLINE-95095951; Pubmed-8002605;
                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25951 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U41011; AAA82289.1; WORMPEP; D2024.7; CE04296. Hypothetical protein. SEQUENCE 225 AA; 25951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 25.2%
Best Local Similarity 33.0%
Matches 74; Conservative
                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                             Caenorhabditis elegans.
                                                            (Rel. 36, (Rel. 36,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pseudomonas putida.
                                                                                                                                                                                                           SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Plasmid NAH7
                                                              15-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pseudomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eaton R.W.;
                              YYS7_CAEEL
Q18973;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAHD_PSEPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           051948;
                                                                                                                              D2024
                YYS7_CAEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAHD_PSEPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAHD.
RESULT
                                 ū
                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66 MANDLKLLRHHLQIPIHFPKDFLSVMLEKG---SLSAMRFLTAVNLEHPEMLEKASRELW 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123 MRVWSRNEDITEPQSILAA--AEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 111 NAVW--GEGIAPDLESLPALVSEKLGWDRSAFEHFLSSNAATE---RYDEQTHAAIERKV 165
                                                                                                                                                                                                                                                                                                                                                                                              recombinant bacteria to study ring cleavage of 1.2-
dihydroxynaphthalene and subsequent reactions.";
J. Bacteriol. 174:554-7554(1992).
-!- FUNCTION: CATALYZES THE ISOMENIZATION OF 2-HYDROXYCHROMENE-2-
CARBOXYLATE (HCCA) TO TRANS-O-HYDROXYBENZYLIDENEPYROVATE (THBPA).
THE OPTIMUM PH FOR THE ENZYME IS 10. THE REACTION IS REVERSIBLE.
-!- CORCIONS: GLUTATHIONE SEEMS TO STABILIZE THE ENZYME, WHICH LOSES ACTIVITY RAPIDLY IN THE ABSENCE OF THIS COMPOUND.
-!- PATHWAY: UPPER NAPTHALENE CATABOLIC PATHWAY WHICH INVOLVES CONVERSION OF NAPTHALENE TO SALICYLATE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY 65
                                  sequence of the DNA encoding 2-hydroxychromene-2-carboxylate
isomerase and trans-o-hydroxybenzylidenepyruvate hydratase-aldolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 VDFYFDFLSPFSYLANQRLSKLAQDYGLTIRYNAIDLARVKIAIGNVGPSNRDLKVKLDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : ||: ||: ||: |: | :: |
LKVDLQRWAQLYGIPLVFPANYNSRRMIGFYYSGAEAQAAAYVNV------VF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTRA_ECOLI STANDARD; PRT; 962 AA.
P05458; P78106;
01-NOV-1988 (Rel. 09, Created)
01-NOV-1988 (Rel. 09, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
PROTEASE III PRECURSOR (EC 3.4.24.55) (PITRILYSIN) (PROTEASE PI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Finch P.W., Wilson R.E., Brown K., Hickson I.D., Emmerson P.T., "Complete nucleotide sequence of the Escherichia coli ptr gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28;
Organization and evolution of naphthalene catabolic pathways:
                                                                                                                                                                                                                                                                                                                                                               "Bacterial metabolism of naphthalene: construction and use of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8.0%; Score 95; DB 1; Length 203; 24.2%; Pred. No. 0.21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U09057; AAA66358.1; -.
ISOMerase; Plasmid; Aromatic hydrocarbons catabolism.
SEQUENCE 203 AA; 23061 MW; E84B56F21C604945 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FGLPITVAHVDGQTHMLFGSDRMELLAHLLG 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ` PRT;
                                                                                                                                                                                                                                                STRAIN-G7 / ATCC 17485;
MEDLINE-93077433; PubMed-1447127;
Eaton R.W., Chapman P.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-87040734; PubMed-3534791;
                                                                                                          m the NAH7 plasmid.";
Bacteriol. 176:7757-7762(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                     CHARACTERIZATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sscherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTRA OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   166
                                                                                                              From
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AC DATE OF THE STATE OF THE STA
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -I- FUNCTION: ENOOPEPTIDASE THAT DEGRADES SMALL PEPTIDES OF LESS THAN 7 KDA, SUCH AS GLUCAGON AND INSULIN.
-!- CATALYTIC ACTIVITY: PREPERENTIAL CLEAVAGE OF 16-TYR-!-LEU-17 AND 25-PHE-|-TYR-26 BONDS OF OXIDIZED INSULIN B CHAIN, ALSO ACTS ON OTHER SUBSTRATES OF WW LESS THAN 7 KDA SUCH AS INSULIN AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-93277449; PubMed-8099278;
Becker A.B., Roth R.A.;
"Identification of glutamate-169 as the third zinc-binding residue in
proteinase III, a member of the family of insulin-degrading
                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-88005781; PubMed-3308636; Claverie-Martin F., Diaz-Torres M.R., Kushner S.R.; Manalysis of the regulatory region of the protease III (ptr) gene of Escherichia coli K-12.";
                                                    SEQUENCE FROM N.A.
STRAIN-K12 / MG1655.
MEDLINE-97426617; PubMed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBCELLULAR LOCATION: PERIPLASMIC. SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
                                                                                                                                                                                                                                                                                                                                  "Complete nucleotide sequence of the Escherichia coli recB gene."; Nucleic Acids Res. 14:8573-8582(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COFACTOR: REQUIRES DIVALENT CATIONS FOR ACTIVITY. BINDS ZINC
                                                                                                                                                                                                                                                                                       Finch P.W., Storey A., Chapman K.E., Brown K., Hickson I.D., Emmerson P.T.;
                                                                                                                                                                                          "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92237263; PubMed-1570301; Becker A.B., Roth R.A.; "An unusual active site identified in a family of zinc metalloendopeptidases.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Proc. Natl. Acad. Sci. U.S.A. 89:3835-3839(1992).
encoding protease III.";
Nucleic Acids Res. 14:7695-7703(1986).
                                                                                                                                                                                                                                                       SEQUENCE OF 853-962 FROM N.A. MEDLINE-87066729; Pubmed-3537960;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL, U29581; AAB40468.1; --
EMBL, RA000385: AAC75860.1; --
EMBL, X06227; CAA39576.1; --
EMBL, M17095; AAA24436.1; --
EMBL, M17095; AAA24436.1; --
MEROPS; M16.001; --
ECO2DBASE; G095.0; GTH EDITION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INTERPRO; IPR001431; -. PFAM; PF00675; Peptidase_M16; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biochem. J. 292:137-142(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MUTAGENESIS, AND ACTIVE SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MUTAGENESIS, AND ACTIVE SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X04581; CAA28249.1; -.
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-296 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ECOGENE; EG10786; PTRA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INSULINASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3ene 54:185-195(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBUNIT: MONOMER.
                                                                                                                                                       Gregor J., Davis
Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLUCAGON.
                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-K12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             enzymes."
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send-an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                         ю
.,
                                                                                                                               LOSS OF ACTIVITY AND OF ZN-BINDING.
LOSS OF ACTIVITY AND OF ZN-BINDING.
20% LOSS OF ACTIVITY.
LOSS OF ACTIVITY.
NO LOSS OF ACTIVITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         J. Bacteriol. 175:6890-6901(1993).
-!- FUNCTION: CATALYZES THE ISOMERIZATION OF 2-HYDROXYCHROMENE-2-CARBOXILATE (HCCA) TO TRANS-O-HYDROXYBENZYLIDENEPYRUVATE (THEPA).
THE OPTIMUM PH FOR THE BENZYME IS 10. THE REACTION IS REVERSIBLE.
-!- PATHWAY: UPPER NAPTHALENE CATABOLIC PATHWAY WHICH INVOLVES CONVERSION OF NAPTHALENE CASALICYLATE, AND CATABOLISM OF DIBENZOTHLOPHENE (DBT) AND PHENNANTHENE. CONVERTS PHENANTHENE TO 1-HYDROXY-2-NAPTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO OXIDATION OF THE ARRANATIC RING.
-!- MISCELLAMEDUS: DOXH AND DOXJ ENCODE DIFFERENT ENZYMES THAT MAY HAVE INTERCHANGEBELE FUNCTIONS.
                                                                                                                                                                                                                                                                                                                                                                       -----QGLLEKIATPKVKNQ 167
                                                                                                                                                                                                                                                                                                                                                                                                    731 WCRNKDVVVDKKQSVIFEKAGNSTDSALAAVFVPTGYDEYTSSAXSSLLGQIVQPWFYNQ 790
                                                                                                                                                                                                                                                                                                                                       23; Gaps
                 4ydrolase; Metalloprotease; Magnesium; Periplasmic; Zinc; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-94042852; PubMed-8226631;
Denome S.A., Stanley D.C., Olson E.S., Young K.D.;
Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: complete DNA sequence of an upper naphthalene catabolic pathwav.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     791 LR--TEEQLGYAVFAFPMSVGRQWGMGFLLQSNDKQ---PSFLWERYKAFFPTA 839
                                                                                                                                                                                                                                                                                                                                                                                                                                      168 LKETTEAACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPA 221
                                                                                                                                                                                                                                   IIIHYVPA -> HYHSLRPW (IN REF. 4)
                                                                                                                                                                                                                                                                                                      7.6%; Score 89.5; DB 1; Length 962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0-MAY-2000 (Rel. 39, Last sequence update)
0-MAY-2000 (Rel. 39, Last annotation update)
-HYDROXYCHROMENE-2-CARBOXYLATE ISOMERASE (HCCA ISOMERASE).
                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                     0558C68C2F1A0540 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Isomerase; Plasmid; Aromatic hydrocarbons catabolism.
SEQUENCE 212 AA; 24039 WW; 5EF96A619913DB4F CRC64
                                                                                                                                                                                                                                                                                                                                     48;
                                                                                                                                                                                                                                                                                                                  26.3%; Pred. No. 4.4; ive 13; Mismatches
                                                 PROTEASE III.
                                                                                                                                                                                PRT;
                                                                                                                                                   E->0:
                                                                                                                                                                    H->R:
                                                                                                                                                                                                                                                                                                                                                                   126 WSRNEDITEPQSILAAAEKAGMSAEQA--
                                                                ZINC
                                                                                                  ZINC
PROSITE; PS00143; INSULINASE; 1
                                                                                                                                                                                                                                                   107708 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last seq
30-MAY-2000 (Rel. 39, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pseudomonas sp. (strain C18)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M60405; AAA16133.1; -.
                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                98
91
92
                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 30; Conserv
                                                                                                                                                                                                                                                   962 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                               24
88
91
169
88
91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAHD_PSESP
Q52462;
                                                                              ACT_SITE
METAL
                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                MUTAGEN
MUTAGEN
                                                                                                                                 MUTAGEN
                                                                                                                                                   MUTAGEN
                                                                                                                                                                 MUTAGEN
                                                                                                                                                                                    dutagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Plasmid
                                                 CHAIN
                                                                                                                 TETAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAHD_PSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
```

5EF96A619913DB4F CRC64;

SEQUENCE

```
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license greenent (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                   MANDLKLLRHHLQIPIHFPKDFLSVMLEKGSL--SAMRFLTA-VNLEHPEMLEKASRELW 122
                                                                                                                                                                 76 LKVDLORWAELYEIPLVFPANYNSRRMNTGLYYSGAMAQTGAYVNV-----VF 123
                                                                                                                                                                                                    MRVWSRNEDITEPQSILAA--AEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                         MEDLINE-90264409; PubMed-2245170; Fischer-Colbrie R., Gutierrez J., Hsu C.M., Iacangelo A., Eiden L.E.; Sequence analysis, tissue distribution and regulation by cell depolarization, and second messengers of bovine secretogranin II
                                                                8 VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY 65
                                                                                  J. BIOL. Chem. 265:9208-9213(1990).
-!- FUNCTION: SECRETOGRANIN II IS A NEUROENDOCRINE SECRETORY GRANULE.
PROTEIN, WHICH IS THE PRECURSOR FOR BIOLOGICALLY ACTIVE PEPTIDES.
                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- MISCELLANEOUS: BINDS CALCIUM WITH A LOW-AFFINITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoldea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
SECRETGRANIN II.
UNKNOWN ACTIVITY PEPTIDE (PROBABLE).
SULFATATION (BY SIMILARITY).
SDC079F559D83516 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sulfatation; Cleavage on pair of basic residues; Calcium-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1- SUBCELLULAR LOCATION: NEUROENDOCRINE AND ENDOCRINE SECRETORY
                                 28;
7.5%; Score 89; DB 1; Length 212; 24.2%; Pred. No. 0.75; ive 37; Mismatches 95; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        01-FEB-1991 (Rel. 17, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
SECRETOGRANIN II PRECURSOR (SGII) (CHROMOGRANIN C).
                                                                                                                                                                                                                                                                                                                                                                                       613 AA.
                                                                                                                                                                                                                                                                     181 FGLPITVAHVDGQTHMLFGSDRMELLAHLLG 211
                                                                                                                                                                                                                                                                                       179 FGVP-TMFLGD---EMWWGNDRLFMLENAVG 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INTERPO: IPR001990; -.
PFAM: PF01271; Granin; 1.
PROSITE; PS00422; GRANINS_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                     01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27 . PC
30 . PC
613 . SI
213 . UI
150 . SC
70356 . MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; J05468; AAA30760.1; -.
                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           taurus (Bovine)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1
28
31
181
150
150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A35296; A35296
               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SCG2 OR CHGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRANULES
                                 51;
                                                                                                                                                                                                                                                                                                                                                                    SG2_BOVIN
ID SG2_BOVIN
AC P20616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FAMILY
Query Match
Best Local S
Matches 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOD_RES
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                   q
                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                     δ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Bodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham M. M. Cleod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Science 281:375-388(1998).

-1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A])
ADDACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).

-1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA PROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.

-1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
                                                                                                                                                            61 RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                  333 RAIRLFEKPLDP-----VGIYQLIEISRNLQIPPEDLIDMLK-TGEKP--VEP 377
                                                                                                                                                                                                                                                                                                                                                                                                                                           30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
TRNA DELTA(2)-ISOPENTENYLPYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
                                                                             6 RTVELFYDVLSPYSWLGFEILCRYQNIWNI-----NLQLRPSLITGIMKDSGNKPPGLLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TIGR; TPU05),
INTERPRO: IPR002627; -.
FRAM, PF01715; IPPT; 1.
Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding.
NP BIND 27 34 ATP (POTENTIAL).
                                              29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Complete genome sequence of Treponema pallidum, the syphilis
      Length 613;
                                           41; Indels
                                                                                                                                                                                                                                                                                ---KMLSKNGYPKAPGHAVAEALPEGLSVEDILNLLGMESAANPK 445
                                                                                                                                                                                                                                          121 LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLL--EKIATPK 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Treponema pallidum.
Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
      DB 1;
  7.5%; Score 88.5; DE 23.6%; Pred. No. 3.1; ative 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=NICHOLS;
MEDLINE=98332770; PubMed=9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                        30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last segu
30-MAY-2000 (Rel. 39, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AE001238; AAC65611.1; -.
Ouery Match
Best Local Similarity 23.6%
Matches 39; Conservative
                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MIAA OR TP0637.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FIGR; TP0637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFERASE
                                                                                                                                                                                                                                                                                                                                                                                MIAA_TREPA
083644;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               spirochete.
                                                                                                                                                                                                                                                                                                                                                             MIAA_TREPA
                                                                                                                                                                                                                                                                                404
                                                                                                                                                                                                쉽
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                              음
                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
                          80 DVCDPYE--EYNVFRFQQAVYGI----VPSIL-----RAHKVPIIVGGTGLYLD---AV 124
                                                            ---RFLTAVNL- 108
                                                                                      125 LRQYALVPVE-RNOALRASLRGASLSHMRAVYFSLKDSHAVHNKTDLEDPARLMRAIEIA 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        387 PYGIEVFNVVLEP------LWKGIRSHRGKVLSSFLKAVGSMIPLMDPEYAG 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 LYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEML-EKASRELW 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      433 YYTTEAMRIIRREFDSPDDEMKKTILLVLQK--CSAVESIT-----PKFLREEIAPEFF 484
DVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 PRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPR-KG 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                        109 ----EHPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIA 160
                                                                                                                                                                                                                                                                                                                                                           Fung1; Ascomycota; Saccharomycetes; Saccharomycetales;
                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-UUL-1998 (Rel. 36, Last annotation update)
HYPOTHETICAL 110.0 KDA PROTEIN IN MSU1-HASI INTERGENIC REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7.1%; Score 84.5; DB 1; Length 971; 21.7%; Pred. No. 12; tive 31; Mismatches 64; Indels 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
1: 27D26E4252A788E2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 MRVWSRNEDITEP-----QSILAAAEKAGMSAEQAQGLLEKIATP 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : | | : |: | : | : | : | 485 QKEWVRRVALDRPLNKVVTYTTVTLAKKLGCSY----TIDKLLTP 525
                                                                                                                                                                                                                                971 AA
                                                                                                                                                                                                                                                                                                                                         Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                          73 LRHHLQIPIHFPKDFLSVMLEKGSLSAM-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Transmembrane
                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                         Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               110027 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; 249704; CAA89786.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Local Similarity 21.79
Matches 36; Conservative
                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S0004901; YMR288W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     252
320
356
663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein;
                                                                                                                                                                                                                                                                                                                          YMR288W OR YM8021.14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                782
877
905
971 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hypothetical
                                                                                                                                                                                                                                YM8P_YEAST
P49955;
                                                                                                                                                                                                                                                                                                                                                            Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                 MBP_YEAST
                        g
                                                                                      Ω
                                                                                                                     óγ
                                                                                                                                                  Ω
                                                          δ
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
```

RESULT

8

Gaps

45;

Indels

90;

30; Mismatches

Conservative

Best Local Similarity Matches 41; Conserva

Query Match

7.1%; Score 84.5; DB 1; Length 316; 23.3%; Pred. No. 3;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseelsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                   102 FLTAVNLE-----SRELWMRVWSRNE 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----YSDAERAALVKNIVDDWEGVREDIKIRNLRNFYQVEPEFAERVAA-GTGINL 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  258 DLYVQVLDPKDLDNYDFNPLDATKDW--FEDVFPYEHVGTMTLNRNPDNIFAETESVGFN 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 P----PGLLPRKG-----LYMANDLKLLR---HHLQIPIHFPKDFLSVMLEKGSLSAMR 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QTSSINYEPNSYDTEPKENPAYIEPEQEIRGDISGRLVAEKPNNFGHAKEVWKR----- 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                131 DITEPQSILAAAEKAGM----SAEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGLPITV 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9 ELFYDVLSP-----YSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNK 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROXIMAL HEME LIGAND (BY SIMILARITY). F0A3251469681EBB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.0%; Score 83.5; DB 1; Length 488; 21.3%; Pred. No. 6.3; Live 37; Mismatches 76; Indels 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
ACT_SITE
55 55 BY SIMILARITY.
ACT_SITE 128 BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
SEQUENCE 488 AA; 55869 MW; F0A3251469681EBB CRC64;
                                                                                                                                                                                                                                                      Bacteria; Firmicutes; Bacillus/Clostridium group;
                                                           01-MAR-1992 (Rel. 21, Created)
01-MAR-1992 (Rel. 21, Last sequence update)
01-NOV-1995 (Rel. 32, Last annotation update)
CATALASE (EC 1.11.1.6).
                                                                                                                                                                                                                                                                                   Bacillus/Staphylococcus group; Listeria
                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-91317736; PubMed-1860824;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PFAM: PF00199; catalase; 1.
PRINTS; PR00067; CATALASE.
PROSITE; PS00437; CATALASE.
1; PROSITE; PS00438; CATALASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M75944; AAB53655.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52; Conservative
   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1179; 11PH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 52; Conserv
                                                                                                                                                                                                                        Listeria seeligeri.
                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P21179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A-HV 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A40367
CATA_LISSE
P24168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INTERPRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
   HIDDER REPAREMENT OF THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

481 AEHV 484

```
RA MEDLINE=ZALISHOUS; PURDREGALU/J134;
RA MEDLINE=ZALISHOUS; PURDREGALU/J134;
RA Amanatides P.G., Scherer S.E., ILP.W., Hoskins R.A., Galle R.F.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Baxter E.G., Hellzel R.G., Clamphe M., Pfelffer B.D.,
RA Brandon R.C., Baxter E.G., Hellz G., Clamphe M., Pfelffer B.D.,
RA Abril J.F., Apdrava P.R., Malson C.R., MIKIOS G.L.G.,
Ballew R.M., Basu A. Baxendle J., Bayrektarcglu L., Beasley E.M.,
Beeson K.Y. Benos P.V., Berman B.P., Bhandari D., Bolthakov S.,
Beeson K.Y. Benos P.V., Berman B.P., Bhandari D., Bolthakov S.,
Beeson K.Y. Doup L.E., Davies M., Brokstein P., Cadieu E., Center A., Chadra I.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Contiler P., Caniey S., Dahlke C., Davenport L.B., Davies P.,
Burtis K.C., Busam D.A., Buller H., Gadieu E., Center A., Chadra I.,
RA Gherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA Borkova D., Boltham A.E., Garg N.S., Gelbart W.M., Glasser K.,
Aloston K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
Alali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Jali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Jali M., Mattel B., McIntosh T.G., McLeod M.P., McPherson D.L.,
Rollazzolo M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y., Rage M.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo M., Pittman G.S., Pen S., Pollard J., Purl Y.,
Ralazzolo
                                                                                                                                                                                                                                       Pterygota, Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila
                                                                                                                                                                                                                                                                                                                         STRAIN-OREGON-R, AND CANTON-S; TISSUE-OVARY;
MEDLINE-90231469; Pubmed-1691829;
Endow S.A., Henikoff S., Soler-Wiedziela L.;
"Mediation of melotic and early mitotic chromosome segregation in Drosophila by a protein related to kinesin.";
Nature 345:81-83(1990).
                                                                                                                                                                                              Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                McDonald H.B., Goldstein L.S.B.; "Identification and characterization of a gene encoding a
                                                                                                             Last sequence update)
Last annotation update)
                                           700 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         kinesin-like protein in Drosophila.";
Cell 61:991-1000(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=90275618; PubMed=2140958;
                                        NCD_DROME STANDARD; F 120480; Q9VAG8; 17, Created) 01-FEB-1991 (Rel. 17, Last sequ
                                                                                                                          01-OCT-2000 (Rel. 40, Last an
CLARET SEGREGATIONAL PROTEIN.
NCD OR CA(ND).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 16-700 FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOTOR DIRECTIONALITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-BERKELEY;
KESULT 11
```

Page

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FUNCTION: NCD IS REQUIRED FOR NORMAL CHROMOSOMAL SEGREGATION IN MEIOSIS, IN FEMALES, AND IN EARLY MITOTIC DIVISIONS OF THE EMBRYO. THE NCD MOTOR ACTIVITY IS DIRECTED TOWARD THE MICROTUBULE'S MINUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRINTS; PROO380; KINESINHEAVY.
PROSITE; PSO0411; KINESINLMOTOR_DOMAIN1; 1.
PROSITE; PSO067; KINESINLMOTOR_DOMAIN2; 1.
MOTOR_DOMAIN2; 1.
MOTOR_DOMAIN; Cell division; Microtubules; ATP-binding; Coiled coil;
Melosis; Mitosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
ATP (BY SIMILARITY).
V->F: IN NCD(D); REDUCES MOTOR VELOCITY.
S -> N (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                          CHÁRACTERIZATION OF MUTANT ALLELE NCD(D).
MEDLINE=96283629; PubMed=8670831;
Moore J.D., Song H., Endow S.A.;
M point mutation in the microtubule binding region of the Ncd motor protein reduces motor velocity.";
EMBO J. 15:3306-3314(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 335-700.
MEDINE=96195067; PubMed=8666780;
Sablin E.P., Kull F.J., Cooke N., Vale R.D., Fletterick R.J.;
"Crystal structure of the motor domain of the kinesin-related motor
               Walker R.A., Salmon E.D., Endow S.A.; "The Drosophila claret segregation protein is a minus-end directed motor molecule.";
                                                                                                                                           Lockhart A., Cross R.A.; "Origins of reversed directionality in the ncd molecular motor.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MISCELLANEOUS: NCD(D) IS A MUTANT ALLELE THAT SHOWS ABNORMAL CHROMOSOMAL SEGREGATION.
SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
                                                                                                                                                                                                                                                                                           non-disjunctional on chromosome segregation in Drosophila.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADE043CBCE7FD561 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COILED COIL (POTENTIAL)
                                                                                                                                                                                                                                                     Komma D.J., Horne A.S., Endow S.A.;
"Separation of meiotic and mitotic effects of claret
                                                                                                                                                                                                                 MUTANT ALLELE NCD(D).
MEDLINE-91122049; Pubmed-1825056;
 MEDLINE-91043032; PubMed-2146510;
                                                                                                                          MEDLINE-94155838; PubMed-8112290;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ₩,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL, X52814; CAA36998.1; --: EMBL, M33932, AAA28716.1; --: EMBL, AE003771; AAF56942.1; --: EMBL, X57475; CAA40713.1; --: EMBL, X57475; CAA40713.1; --:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; A35624; A35624.
PIR; S09748; S09748.
HSSP; P17119; 3KAR.
FLYBASE; EBGN002924; ncd.
INTERPRO; IPR001752; -.
PFAM; PF00225; kinesin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77473
                                                                                                                                                                               EMBO J. 13:751-757(1994).
                                                                                                                                                                                                                                                                                                        EMBO J. 10:419-424(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nature 380:555-559(1996)
                                                                   Nature 347:780-782(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          434
556
697
700 AA;
                                                                                                       CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NP_BIND
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                               12
                                                                                                                                                                                                                                                                                                                                                               |:: ||::| |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:
                                                                                                                            53 NKPPGLLPRKGLYMAN---DLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLE 109
                                                                                                                                                                                                                                                                 NK-----NDIYVSNITEETVLDPNHLRHLMHTAK-----MNRATAS-----TAGN-- 547
                                                                                                                                                                                                                                                                                                                               110 HPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMS------AEQAQGLLEK 158
                               73; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Métazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MGPLPRTVELFYDVLSPYSWLG--FEILCRYQNIWNINLQLRPSLITGIMKD-----SG 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- SIMILARITY: BELONGS TO THE CHROMOGRANIN / SECRETOGRANIN PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-89098327; PubMed-3211750;
Gerdes H.-H., Philipps E., Huttner W.B.;
"The primary structure of rat secretogranin II deduced from a CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PFAM; PF01271; Granin; 1.
PROSITE; PS00422; GRANINS_1; 1.
Sulfatation; Cleavage on pair of basic residues; Calcium-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [2]
SEQUENCE FROM N.A.
MEDLINE-93309708; PubMed-8321414;
MEDLINE-93.09708; PubMed-8321414;
Kakar S.S., Wei N., Mulchahey J.J., Leboeuf R.D., Neill J.D.;
"Regulation of expression of secretogranin II mRNA in female rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECRETOGRANIN II.
UNKNOWN ACTIVITY PEPTIDE (PROBABLE).
                                                                                                                                                                                                                                                                                                                                                                                                                                               159 IATPKVKNQLKETTEAACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLL 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- MISCELLANEOUS: BINDS CALCIUM WITH A LOW-AFFINITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <u>:</u>
                               70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
SECRETOGRANIN II PRECURSOR (SGII) (CHROMOGRANIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               619 AA
                               37; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nucleic Acids Res. 16:11811-11811(1988).
   Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X13618; CAA31950.1; -. EMBL; M93669; AAA42135.1; -.
22.4%;
                               52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30
619
216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SCG2 OR CHGC OR SCG-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR001990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S02180; S02180
   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31
184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GRANULES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INTERPRO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIĞNAL
                                                                                                                                                                                                                                                                    508
                               Matches
                                                                                          1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
SG2_RAT
                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                              q
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                   œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N PYROPHOSPHATE + DNA(N).
-1- MISCELLANEOUS: THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER
-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-87106854; PubMed-3803925;
Shu L., Hong J.S., Wel Y.-F., Engler J.A.;
"Nucleotide sequence of the genes encoded in early region 2b of human
                                                                                                                                                                                                 73 LRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDI 132
                                                                                                                                                                                                                        13 DVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYMANDLKL 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human adenovirus type 12.
Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.; "Nucleotide sequence of human adenovirus type 12 DNA: comparative
                                                                                                   51;
                                                                  DB 1; Length 619;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA-directed DNA polymerase; DNA replication;
SULFATATION (BY SIMILARITY).
27CB75B4F25A38D1 CRC64;
                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              adenovirus typė 12.";
Gene 46:187-195(1986)
-!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
                                                                                                                                                                                                                                                                  133 TE----PQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTE 173
                                                                                                                                                                                                                                                                                     188 VEEQYTPQS-LATLESVFQE------LGKLTGP--SNQKREKVD 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R -> S (IN REF. 2).
V -> L (IN REF. 2).
LQ -> YN (IN REF. 2).
S -> T (IN REF. 2).
L -> F (IN REF. 2).
                                                                                                 57;
                                                                                                                                                                                                                                                                                                                                                                                                             01-JAN-1988 (Rel. 06, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
                                                               ; Score 81.5; E
; Pred. No. 13;
16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRINTS; PR00106; DNAPOLB.
PROSITE; PS00116; DNA_POLYMERASE_B; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X73487; CAA51882.1; -.
EMBL; M14785; AAA42478.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. MEDLINE=94076430; PubMed=8254750;
               71031 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA POLYMERASE (EC.2.7.7.7).
                                                                                24.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 functional analysis.";
J. Virol. 68:379-389(1994).
                                                        Query Match
Best Local Similarity 24.8%
Matches 41; Conservative
                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32
162
182
461
575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INTERPRO; IPR002064;
               619 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A25770; DJAD12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              833933;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transferase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA-binding
                                                                                                                                                                                                                                                                                                                                                                                DPOL_ADE12
P06538;
 MOD_RES
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                 DPOL_ADE12
                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
                                                                                                                                                                                                                                                                                               ------DDNLVKHALQLALELGWDQWEKDSTTFYCLTPEKMKV-GQQFRTYRN--- 218
                                                                                                                                                                                                                                                                                                                                     111 REMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKE 170
                                                                                                                                                                                                                     133 IGSHPRTERLFITYDV-ETYTWMG------AFGKQLVPFML--VMKLSG----- 172
                                                                                                                                                                                                                                                           59 LPRKGLYMANDLKLLRHHLQIPI-----HFPKD---FLSVMLEKGSLSAMRFLTAVNLEH 110
                                                                                                                                                                                                                                                                                                                                                                        219 - RLQTSLATDLWMTFLQKNPHLSQ-----WAQEENGLVALEDLSYEDLKRAPAIKGEPR- 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67; Gaps
                                                                                                                                                                               1 MGPLPRTVELF -- YDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               38; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7 TVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLPRKGLYM 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
HYPOTHETICAL TRNA/RRNA METHYLTRANSFERASE SLR1673 (EC 2.1.1.-).
                                                                                                  DB 1; Length 1061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 6.8%; Score 81; DB 1; Length 274; Best Local Similarity 21.1%; Pred. No. 5.1; Matches 35; Conservative 32; Mismatches 61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                171 TTEAACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------FVELYIVG--HNINGFDEIVLAAQVINNRLDVPGP 304
                                                                                                                                        Indels
                                          33FBA89C33065C08 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INTERPRO, IPR001537; -. PFAM; PF00588; SpoU_methylase; 1. Hypothetical protein; Transferase; Methyltransferase. SEQUENCE 274 AA; 29701 MW; A4176C1061CAAE88 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Cyanobacteria; Chroococcales; Synechocystis
                                                                                                                                    77;
S -> T (IN REF. 2).
K -> M (IN REF. 2).
MW; 33FBA89C33065C0
                                                                                                6.9%; Score 81.5; D
21.8%; Pred. No. 25;
1ve 35; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synechocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-97061201; Pubmed-8905231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Rel. 37, Created)
(Rel. 37, Last sequ
(Rel. 37, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D90913; BAA18355.1; -.
  892 892
1030 1030
1061 AA; 121727
                                                                                                                    21.8%;
                                                                                                  Query Match 6.9%
Best Local Similarity 21.8%
Matches 50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-DEC-1998
15-DEC-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YG73_SYNY3
P74261:
    CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SLR1673
                                                                                                                                                                                                                                                                                                 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                      272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KG73_SYNY3
    FT
                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                          g
```

σ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                   67 ANDL-----KLLRHHLQIPIHFPK---DFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKA 117
                                  152 TADCVDPTSPKVLRSSAGSSLLLPQQQLQSLPPLLEKFHTQGLQLIATV-----PQA 203
                                                                                                                                                                                                                                 01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE ATPASE SUBUNIT (EC 4.1.1.21)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRK-GLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSL-----SAMRFLTAVNLEH 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      158 PAILEGEVEFEREV------SVIAARDRSGNVAIFDLAENVHKDGILATST 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              203 VPAAISVQTAEAARTAAEKLLHALDYVGVLGLEFFVLKDGTLLANEFAPRVHNS-GHWTE 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----TPKVKNQLKETTE 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PLPRTVELFYDVLS---PYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLL 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47 PLPTSSPMNLKMCQSAPPTSWLKRRLFCPRPPHWK-----SLRTASQKSSFSTKAALK 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 AACRYGAF-----GLPI--TVAHVDGQTHMLFGSD------RMELLAHLLGEK 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BY SIMILARITY).

CATALYTIC ACTIVITY: 1-(5-PHOSPHORIBOSYL)-5-AMINO-4-IMIDAZOLE-CARDXYLATE -1-(5-PHOSPHORIBOSYL)-5-AMINOIMIDAZOLE + CO(2).

PATHWAY: SIXTH STEP IN DE NOVO PURINE BIOSYNTHESIS.

SUBUNIT: HOMODIMER (BY SIMILARITY).

SIMILARITY: BELONGS TO THE PURK / PURT FAMILY.
                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96; Indels 113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.8%; Score 80.5; DB 1; Length 339;
                                                                      SRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPK 163
                                                                                            204 TQTLW-----EIDFORPTIVIFGSEGGGLSAPVLELTTHQVAIPQ 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Purine biosynthesis; Lyase; Decarboxylase.
SEQUENCE 339 AA; 37166 MW; 951B0DC5E637324E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      111 PEMLE---KASRELWMRVWSRNEDITEPQSILAAAEKAG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            147 -----MSAEQAQGLLEKIA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity 21.7%; Pred. No. 7.3; 65; Conservative 25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; U10241; AAA57003.1; -.
                                                                                                                                                                                                                                                                                                         (AIR CARBOXYLASE) (AIRC).
                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                Brucellaceae; Brucella.
                                                                                                                                                                                                                                                                                                                                             Brucella melitensis.
                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-16M;
                                                                                                                                                                                                  PURK_BRUME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local S
Matches 65
                                                                                                                                                             RESULT 15
PURK_BRUME
                                                                      118
                                  Op
                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
```

Db 262 AACAISQFEQHIRAVAGLPLGNTDRHSDCVMENLIGDDIEKVPAILCEKNAVLHLYGKK 320

Search completed: April 3, 2001, 11:49:17 Job time: 632 sec us-09-441-723-1.rsp

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

April 3, 2001, 11:38:20 ; Search time 25.94 Seconds (without alignments) 1021.164 Million cell updates/sec

US-09-441-723-1 1185 1 MGPLPRTVELFYDVLSPYSW.....AHLLGEKWMGPIPPAVNARL 226 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

374700 seqs, 117207915 residues Searched:

374700 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SPTREMBL_15:* Database

1: sp_archea:*
2: sp_bacteria:*
3: sp_fung1:*
5: sp_uman:*
6: sp_mammal:*
7: sp_mhc:* sp_virus:*
sp_vertebrate:*
sp_unclassified:* sp_organelle:* sp_phage:* sp_plant:* sp_rodent:* 5: 6: 7: 8: 9: 10:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

20 82 6.9 203 2 007298 21 80.5 6.9 296 2 052799 22 80.6 6.8 10 055098 24 80 6.8 413 10 095095 25 80 6.8 402 1 055098 26 80 6.8 221 5 009355 26 80 6.8 821 2 030926 27 8.5 6.6 493 10 095814 29 78.5 6.6 693 5 099891 31 78.5 6.6 693 5 099891 32 78 6.6 1072 6 077690 33 78 6.6 1072 6 077690 34 77 5 6.5 754 11 092018 41 77.5 6.5 1166 4 099783 43 77 6.5 659 2 094873 44 77 6.5 659 2 094874	O07298 pseudomonas O52799 rhizobium 1 O55098 mus musculu O91up2 arabidopsis	CEDO	Q9rsb4 deinococcus Q9svj3 arabidopsis Q9vs91 drosophila Q77690 bos taurus	O29435 archaeoglob O9yfi8 aeropyrum p O84200 chlamydia t O66234 escherichia O27560 methanobact	Q92277 mus musculu Q9xbn8 bacillus st Q920r0 mus musculu Q91114 mus musculu Q952r3 homo sapien	Q9uya3 pyrococcus Q9v977 drosophila Q9z8d0 chlamydia p Q9jry4 chlamydia p
82 6.9 203 81 6.8 296 80.5 6.8 413 80 6.8 221 80 6.8 704 80 6.8 704 80 6.8 821 79 6.7 230 78.5 6.6 693 78.5 6.6 693 78.5 6.6 1072 78 6.6 338 78 6.6 338 78 6.6 348 77 5 6.5 1166 77 5 6.5 1166 77 6.5 6.5 1166	007298 052799 055098 09LUP2	0000	0 00	029435 Q9YFI8 084200 066234 027560	Q92277 Q9XBN8 Q9Z0R0 Q9JJJ4 Q9P2R3	Q9UYA3 Q9V977 Q9Z8D0 Q9JRY4
82 6.9 81 6.8 80.5 6.9 80 6.8 80 6.8 80 6.8 80 6.8 77 78 5 6.6 77 5 6.5 77 6.5	2 11 10			44004	12114	
80 80.5 80.5 80.5 80 80 80 80 74 74 77 77 77 77 77 77 77 77 77 77 77	203 296 966 413	221 402 704 821	230 493 693 1072	248 313 338 704	1479 482 754 754 1166	24 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
	0 0 0 0 0		0 0 0 0 0 0 0	00000 00000	0 0 0 0 0 0 0 0 0 0	9 9 9 9 9
01122222222222222222222222222222222222	0)	78.5 78.5 78.5	78 78 78 78	78 77.5 77.5 77.5	7.7 7.7 7.7
	20 21 22 23	25 26 27	3 3 5 8 3 3 5 8	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	33 38 44 04 11	4444 5444 5443

ALIGNMENTS

19Y203: 1-NOV-1999 (TrEMBLrel. 12, Created) 1-NOV-1999 (TrEMBLrel. 12, Last sequence update 1-NOV-1999 (TrEMBLrel. 12, Last sequence update 1-NOV-1999 (TrEMBLrel. 12, Last sequence update 1-NOV-1999 (TrEMBLrel. 12, Last annotation update 1-NOV-1999 (TrEMBLrel. 12, Last annotation update 1-NOV-1999 (TrEMBLrel. 12, Last annotation update 2-NOV-1999 (TrEMBLrel. 12, Last annotation update 3-NOV-1999 (TrEMBLrel. 13, NOV-1999) 13 Fundalia; Eutheria; Primates; Catarrhini; Hominian EQUENCE FROM N.A. ao M., Ye M., Zhang O., Zhou J., Wu J., Shen Y. Wa B., Fu G., Chen S., Chen Z.; Human rGSTK1-1 homolog gene."; ubmitted (JUN-1998) to the EMBL/GenBank/DDBJ damblited (JUN-1)				()			ta; Euteleostomi;	dae; Homo.				, Kan L., He K.,			tabases.		CRC64;	Length 226;	Indels 0; Gaps	ITGIMKDSGNKPPGLLP 60	ITGIMKDSGNKPPGLLP 60	TAVNLEHPEMLEKASRE 120	TAVNLEHPEMLEKASRE 120	VKNQLKETTEAACRYGA 180	VKNQLKETTEAACRYGA 180	IARL 226	
99 y 20 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	226		eated)	st sequence appare	NIT 13 HOMOLOG.		Craniata; Vertebra	Catarrhini; Homini				:			BL/GenBank/DDBJ da			Score 1185; DB 4;	Pred. No. 4.9e-90; Mismatches 0;	LCRYQNIWNINLQLRPSL	LCRYQNIWNINLQLRPSL	DFLSVMLEKGSLSAMRFL	DFLSVMLEKGSLSAMRFL	GMSAEQAQGLLEKIATPK	GMSAEQAQGLLEKIATPK	LAHLLGEKWMGPIPPAVN	
99 y 20 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		,	(TrEMBLrel.	(Tremberel.	S-TRANSFERAS	(Human).	etazoa; Chordata;	Primates;				., Zhang Q., Zhou	, Chen S., Chen Z.	1-1 homolog gene."	UN-1998) to the EM	. 'T: 50605000 '/6		100.08;	100.0%; vative 0;	VELFYDVLSPYSWLGFEI	VELFYDVLSPYSWLGFEI	NDLKLLRHHLQIPIHFPK	NDLKLLRHHLQIPIHFPK	RNEDITEPQSILAAAEKA	RNEDITEPOSILAAAEKA	AHVDGQTHMLFGSDRMEL	
	9Y2Q3				GLITTATHIONE	Homo sapiens	Eukaryota; M	Mammalia; Eu	NCBI_TaxID=9	[1]	SEQUENCE FRO	Mao M., Ye M	Gu B., Fu G.	"Human rGSTK	Submitted (J	Transforace	SEQUENCE 2	Query Match	Best Local Simi Matches 226;								_

ö

64

105

```
65 YMANDLKLLRHHLQIPIHFPKDFLSV------MLEKGSLSAMRFLTAVNLEHPEMLEK 116
                       2 RTLDFYFDYRSPYSFLALSQVRK----MDVEIAFHPLEIGDLMKQVGNVPTSITCAPKGR 57
RTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL-LPRKGL
                                                                 Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   195 AA; 22064 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01,
01,
07,
                                                                                                                                                                                                                                                                (TrEMBLrel. 12,
                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; U65001; AAD45416.1; Isomerase.
                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                              Submitted (JUL-1999) to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel.
01-NOV-1996 (TrEMBLrel.
01-AUG-1998 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=28214;
                                                                                                                                                                                                                                                                                                                                Sphingomonas sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COSMID T07E3.
                                                                                                                                                                                                                                                                                                                                                      Sphingomonas
                                                                                                                                                                                                                                                                 01-NOV-1999
                                                                                                                                                                                                                                                                                        01-NOV-1999
                                                                                                                                                                                                                                                                             01-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-BN6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                   Keck A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          T07E3.3
                                                                                                                                                                                                                                        09x907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                022312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ŋ
                                                                                                                                                                                                                                                                                                                  NSAD.
                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     022312
                        g
                                                 ò
                                                                         qq
                                                                                                 ò
                                                                                                                         a
                                                                                                                                                   ð
                                                                                                                                                                         g
                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Brito B., Palacios J., Imperial J., Ruiz-Argueso T.; Identification of a gene for a chemoreceptor of the methyl-accepting type in the symbiotic plasmid of Rhizobium leguminosarum bv. viciae UPM791.";
                                                                                                                                                                                                                                                                                                                                                                                        RKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                    LWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                       1 MGPLPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGLLP 60
                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
                                                                                                                                                                                                                                                                                                                ó:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 194;
                                                                                                                                                                                                                                                                                    98.6%; Score 1169; DB 4; Length 226; 99.1%; Pred. No. 1.2e-94; Live 0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                              Zhao Z., Huang X., Li N., Zhu X., Cao X.;
A novel gene from human dendritic cell.";
Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
EMBL, AF068287; AAF65506.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FGLPITVAHVDGQTHMLFGSDRMELLAHLLGEKWMGPIPPAVNARL 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
PUTATIVE GLYCEROL-3-PHOSPHATE TRANSPORT PROTEIN (UGPC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Biochim, Biophys. Acta 1308:7-11(1996).
EMBL; U23040; AAC44312.1; -.
SEQUENCE 194 AA; 21153 MW; 0421C131B37BA435 CRC64;
                                                                                                                                                                                                                                           226 AA; 25586 MW; D3FDB561533B5A65 CRC64;
                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13.9%; Score 164.5; DB 2; 25.2%; Pred. No. 7.3e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34; Mismatches
                                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-96328256; PubMed-8765742;
                                                                     01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15,
                                                                                                                                                                                                                                                                                                             Matches 224; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rhizobium leguminosarum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rhizobiaceae; Rhizobium.
NCBI_TaxID=384;
                                                                                                                        Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 54; Conserve
                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                           HDCMD47P
                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                 Best Local
                                             Q9P1S4
Q9P1S4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            052782
                      ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      m
                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                   Q9P1S4
                                               ò
                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
```

```
117 ASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAAC 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RKGLYMANDLKLLRHHLQIPIHFPKDF-----LSVMLEKGSLSAMRFLTAVNLEHPEML 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                115 EKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEA 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       112 -----YHRIWGIGIDPRDQNELRGCAEDVGLDAD---ALCEFVRSPAGQGEYRKARTQ 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 LPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPG---LLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MIKTIDFYFDFISPFSYLAQVKLPDLARRTGCVIEYRPIDIPEAKIAAGNYGPSNREVVP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 12.6%; Score 149; DB 2; Length 19
Best Local Similarity 23.1%; Pred. No. 1.7e-05;
Matches 50; Conservative 43; Mismatches 91; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7E6FBAC96B3C6316 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
2-HYDROXYCHROMENE-2-CARBOXYLATE ISOMERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              175 ACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLL 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161 SRGVFGAPTLFV----GDEMFFGNDRLHFMQGHL 190
                                                                                                        177 RYGAFGLPITVAHVDGQTHMLFGSDRMELL-AHL 209
```

7;

```
Querr
Best Local Simi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Burkholderia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                     Query Match
             SECUENCE
                                                                                                                                                                                                                                                                                                                                                                          092HH4;
                                                                                                                                                                                                                                                                                                                                                            Q92HH4
                                                                                                                                                                                                                                                                                                                                   7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                ò
  SE
                                                                                                                                                                                   g
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                               Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 VHNSRLWKTKIEVPEQAESEETLSEMGILQKIDERGKKLIGCERVVPPVDWKNTYKTAVA 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94 KGSLSAMRFLTAVNLEHPEMLEKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQ 153
                                                                                                                                                                                                                                                                                                                                                                                                                                            MGPLPRTVELFYDVLSPYSW-------LGFEILCRYQ-----NIWNINLO 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                           39 LRPSLI--TGIMKDSGNKPPGLLPRKGLYMAND---LKLLRHHLQIPIHFPKDFLSVMLE 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Romine M.F., Stillwell L.C., Wong K.-K., Thurston S.J., Sisk E.C., Sensen C.W., Gaasterland T., Saffer J.D., Fredrickson J.K.; "Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromaticivorans strain F199.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
                                                                                                                                                                                                                                                                                                                                                                                                                 29;
                                                                                                                                                                                                                                                                                                                                                                                        Length 312;
                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180 DIVARLSSPENRSILHKNCKEAVDFKLTEAPGLILLTDEGDTIKIF 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases EMBL; AF079317; AAD04010.1; -.
                                                                                                                                                                                                                                                     Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                         Waterston R.; Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       154 GLLEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVD-GQTHMLF 198
                                                                                                                                                                                                                                                                                                                                  EMBL; U13643; AAA21082.1; -.
SEQUENCE 312 AA; 35436 MW; 0FB8561F57687120 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
2-HYDROXYCHROMENE-2-CARBOXYLATE ISOMERASE.
                                                                                                                                                                                                                                                                                                                                                                                     12.3%; Score 145.5; DB 5; 22.6%; Pred. No. 6.3e-05; tive 46; Mismatches 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sphingomonas aromaticivorans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1998 (TrEMBLrel. 08, 01-NOV-1998 (TrEMBLrel. 08, 01-NOV-1998 (TrEMBLrel. 08)
                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 22.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                              SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-48935;
NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-F199;
                                                                                                                                                                                                                                                                                                                                                                                        Local Similar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sphingomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Plasmid pNL1
                                                                                                                                                                                                                                          Favello A.;
                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local S
Matches 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 085994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAHD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

å

ô

```
strain RP007 constitute a divergent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 VWSRNEDITEPQSILAAAEKAGMS-AEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGL 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           117 GGAPDDD-----AALRSIAVSMGWDAADFLRFLDSSEAETAYNESTLEAISAGVFGV 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 -KIKVMKADLERWAERYGVPLTFPASFACADWNCAVLFAREHGKAEAFVT----- 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              65 YMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASRELWMR 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 RKGLYMANDLKLLRHHLQIPIHFPKDF-----LSVMLEKGSLSAMRFLTAVNLEHPEML 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EKASRELWMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEA 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                110 -----DAYRRIWGOGIDPGDRNELAACAIAAGL---DPAALIAFVESPAGONEYRKARSO 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 TIDFFFDFLSPYAYLARHRLTQVAALHGCAIAYKPIDLARAKLAIGNTGPANRDMPVKLA 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7 TVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGL 64
                                                                                                                                                                                                                                                                                                                                                      32; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 LPRTVELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPG---LLP 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : ||:: ::| :||:||:|| | : :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Laurie A.D., Lloyd-Jones G.;
"The phn genes of Burkholderia sp. strain RP007 constitute a d.
"The phn genes of Burkholderia aromatic hydrocarbon catabolism.";
gene cluster for polycyclic aromatic hydrocarbon catabolism.";
J. Bacteriol. 181:531-540(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 196;
                                                                                                                                                                                                                                    Length 197;
                                                                                                                                                                                                                              11.8%; Score 140; DB 2; Length 19
23.1%; Pred. No. 0.0001;
tive 40; Mismatches 94; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       196 AA; 21908 MW; 58716E9039BEA76B CRC64;
                                                         21998 MW; COCD74B7106C07CE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11.4%; Score 135; DB 2; L6
24.0%; Pred. No. 0.00028;
tive 35; Mismatches 103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     175 ACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHLL 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTMAVGRD----MWWGNDRIDFLETHLG 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          184 PITVAHVDGQTHMLFGSDRMELLAHLLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=99102216; PubMed=9882667;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10), 01-MAY-1999 (TrEMBLrel. 10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF061751; AAD09871.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 50; Conservative
                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Burkholderia sp. RP007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Plasmid.
197 AA;
                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=83784;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ISOMERASE PHND.
```

5

. Θ

```
7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KLTHLSRDLQRWAQRYGTPLKFPPSFDSRRLNTGF----FYAAGEAREAEYVRRA---- 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 KGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASREL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 WMRVWSRNEDITEPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                   2 LSFYFDFVSPFSYLASIRLPEIVQRY----GISVSYKPIDIACAKRAIGNVGPSNRDMPV 57
                                                                                                                                                                                                                                                                                                                                                                                        26; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                  8 VELFYDVLSPYSWLGF----EILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPR 61
                                                                                                                                                                                                             Kiyohara H., Tabata Y., Takizawa N.;
"A phenanthrene degradative gene cluster in Alcaligenes faecalis
AFK2.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Fuenmayor S.L., Wild M., Boyes A.L., Williams P.A.;
"A gene cluster encoding steps in conversion of naphthalene
gentisate in Pseudomonas sp. strain U2.";
J. Bacteriol. 180:2522-2530(1998).
                                                                                                            Alcaligenes faecalis.
Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
Alcaligenes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
                                                                                                                                                                                                                                                                                                                                                            Length 196;
                                                                                                                                                                                                                                                                                                                                                                                      99; Indels
                                                                                                                                                                                                                                                                 Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AB024945; BAA76325.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=U2;
Fuenmayor S., Wild M., Boyes A.L., Williams P.A.;
Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF036940; AA012617.1;
                                                                                                                                                                                                                                                                                                               196 AA; 22355 MW; C027579783BEE2E2 CRC64;
                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
2-HYDROXYCHROMENE CARBOXYLATE ISOMERASE.
                                                                                                                                                                                                                                                                                                                                                        Query Match 9.6%; Score 114; DB 2; Best Local Similarity 21.1%; Pred. No. 0.019; Matches 44; Conservative 40; Mismatches 99
               196 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      199 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 GLPITVAHVDGQTHMLFGSDRMELLAHLL 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-98233751; PubMed-9573207;
                                         01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12,
             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pseudomonas sp. U2
                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=70356;
                                                                                                                                                                NCBI_TaxID=511;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Plasmid pWWU2
                                                                                                                                                                                                          STRAIN-AFK2;
                                                                                       ISOMERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ralstonia
                                                                                                                                                                                                                                                                                                              SEQUENCE
            онхмо
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0923X5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        σ
                                                                                                     PHND.
OHXMGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0923X5
             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
```

```
66 MANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRFLTAVNLEHPEMLEKASR---- 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65 LKVDLORWADLYRIPLVFPPNFNSRRVNAG------LYYPAARERAAEYVRLVF 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           120 - ELWMRVWSRNEDITEPQSILA-AAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAAGR 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113 DSAWGKGWALDAD----SLLAEVCDKLNWDLGEFEDFLNSENAAKAYD---EETQAAID 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                125 VWSRNEDIT-EPQSILAAAEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGAFGL 183
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66 MANDLKLLRHHLQIPIHFPKDFLSVMLEKG-SLSAMRFLTAVNLEHPEMLEKASRELWMR 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |: ::| |||:|:| | | : ::| | : ::| | | : ::| | | | : ::| | | | | : ::| | | | : ::| | | : ::| | | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| | : ::| |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bosch R., Lalucat J., Timmis K.N., Moore E.R.B.;
"Complete nucleotide sequence of a chromosomally encoded naphthalene degradation pathway from Pseudomonas stutzeri ANIO and its evolutionary significance.";
Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; AF039533; AAD02142.1;
-. 2589ENCE 238 AA; 26359 MW; 29B972CICCF7C71B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8 VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY 65
                                                                                                                                                                                                                                                                                               8 VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY
                                                                                                                                                                                                                                   34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pseudomonas stutzeri (Pseudomonas perfectomarina).
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 238;
                                                                                                                                                           Length 199;
                                                                                                                                                                                                                               91; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       184 PITVAHVDGQTHMLFGSDRMELLAHLL------GEKWMGPI 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23067 MW; 219466E06AC04E47 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8.2%; Score 97.5; DB 2; 1
23.5%; Pred. No. 0.69;
tive 37; Mismatches 103;
                                                                                                                                            Query Match 8.9%; Score 106; DB 2; Best Local Similarity 24.9%; Pred. No. 0.099; Matches 52; Conservative 32; Mismatches 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2-HYDROXYCHROMENE-2-CARBOXYLATE DEHYDROGENASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 238 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           178 YGAFGLPITVAHVDGQTHMLFGSDRMELL 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 165 RKVFGVP-TVFWDD---QMWWGNDRLFML 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 23.59
Matches 53; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
   Plasmid.
199 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-316;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-AN10;
Isomerase;
                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 791260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   092167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09ZI67
       SOE
                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

ģ

11

RESULT Q9WXQ0

```
051499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
   qq
                                                                                                                                                                                                                      ŏ,
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                     Оp
                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                         óλ
                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                      Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.; Exidence for lateral gene transfer between Archaea and Bacteria from Nature 399:323-329(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DLKLLRHHLQIP----IHFPKDFL-----SVMLEKGSLSAMRFLTAVNLEHP-EMLE 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116 KASRELWMRVWSRNEDITEPQSILAAAEKAGM-----SAEQAQGLLEKIATPKVKNQLK 169
                                                                                                                                                                                                                                                                                                                                                                                                                   28 RYQNIW------LPRKGLYMAN 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Archaea; Buryarchaeota; Methanobacteriales; Methanobacteriaceae; Methanobacterium
                                                                                                                                                                                                                                                                                                                                                                                                   38;
                                                                                                                                                                                                                                                                                                                                                                            7.7%; Score 91; DB 2; Length 581;
19.0%; Pred. No. 8.4;
                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                               581 AA; 66324 MW; 55719E0DCA5D0A05 CRC64;
                 01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (TrEMBLrel. 05, Last sequence update) (TrEMBLrel. 15, Last annotation update)
                                                                                                                                                                                                                                                                                                                PS00211; ABC_TRANSPORTER; UNKNOWN_1.
PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
                                                                                                                                                                                                                                                                                                                                                                                                 67;
581 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  410 AA
                                                                                                                                                                                                                                                                                                                                                                                                 44; Mismatches
                                               ABC TRANSPORTER, ATP-BINDING PROTEIN TM0043.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MTH220.
Methanobacterium thermoautotrophicum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JAN-1998 (TrEMBLrel. 05, Created)
                                                                              Thermotogales; Thermotoga.
PRT;
                                                                                                                     STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; PubMed-10360571;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                           PFAM; PF00005; ABC_tran; 1.
PFAM; PF00664; ABC_membrane; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FLAVOPROTEIN A HOMOLOG (II).
                                                                                                                                                                                                                                                                                                                                                                                     19.0%;
                                                                                                                                                                                                                                                                                                                                                                                     Best_Local Similarity 19.0%
Matches 35; Conservative
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                              INTERPRO; IPRO01140; -.
INTERPRO; IPRO01617; -.
INTERPRO; IPR002106; -.
                                                                    Thermotoga maritima
                                                                                                                                                                                                                                            P13569: 1NBD
                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                       NCBI_TaxID=2336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-2166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           170 ETTE 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ELSK 526
                                                                                                                                                                                                                                                     TM0043;
                                                                                                                                                                                                                                                                                                                         PROSITE; PS
ATP-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JAN-1998
                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                  PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               026322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              523
                                                                                                                                                                                                                                                                                                                                                                                                                                                           69
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55 ----PPG---LLPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVMLEKGSLSAMRF----L 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 MLHWPDSMFTLLEEEGILFSND--AFGOHLCISKRFDKDVPEAVLMD---AAMKFYANLL 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            104 TAVNLEHPEMLEKAS--RELWM------RVWSRNEDITEPQSILAAAEKAGMSA 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               253 DWATGKCRDKAT-IIYDTMHYSTRMLAHAMAEGL--MAADVDVSMHFLHEDERSEIVKNI 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13 DVLSPYSWLGFEILC------RYQNIWNINLQLRPSLITGIMKDSGNK----- 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
MEDLINE-98037514; bubmed-9371463; STRAIN-DELTA H; MEDLINE-98037514; bubmed-9371463; Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBols and Laredge T., Backely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicare R., Wang Y., Wierzbowski J., Glason R., Jiwani N., Caruso A., Bush D., Safer H., Petwell D., Prabhakar S., McDougall S., Shimer G., Goyal A., Pletrovski S., Church G.M., Poniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.; Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "The molecular analysis of an NAH7-type gene cluster, pah, located
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TPLS---PLVLRKFSEVKELGLLEKIGMIAPSHGQIW-----TEPLKIIAA----YT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      150 EQAQGLLEKIATPKVKNQLKETTEAACRYGAFGLPITVAHVDGQTHMLFGSDRMELLAHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
Kiyohara H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.5%; Score 89; DB 2; Length 199; 24.2%; Pred. No. 3; tive 37; Mismatches 95; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the chromosome of Pseudomonas aeruginosa PaK1.";
Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
EMBL; D84146; BAA12247.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                 8EAF9D3B363A8BD5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199 AA; 22509 MW; 81C2A90CA56E5462 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 7.6%; Score 90.5; Di
Best Local Similarity 25.1%; Pred. No. 5.8;
Matches 65; Conservative 31; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    199 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                   PFAM; PF00753; lactamase_B; 1. SEQUENCE 410 AA; 45740 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     210 LGEK------ 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          310 LESKAVFIGSPTMFNGPFP 328
                                                                                                                                                                                                                                                                                                                                            EMBL; AE000809; AAB84726.1;
INTERPRO; IPR001279; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01,
01,
09,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel.
01-NOV-1996 (TrEMBLrel.
01-JAN-1999 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-PAK1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Isomerase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ISOMERASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
```

9

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 188
                                                                                                                       MANDLKLLRHHLQIPIHFPKDFLSVMLEKGSL--SAMRFLTA-VNLEHPEMLEKASRELW 122
                                                                                                                                                                                                                                                    MRVWSRNEDITEPQSILAA - - AEKAGMSAEQAQGLLEKIATPKVKNQLKETTEAACRYGA 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Arabidopsis thallana (Mouse-ear cress).
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnollophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-CV. GMILB;
Mourad G.S., Emerick R.M., Smith A.M.;
"Cloning and Sequencing of a cDNA Encoding an Isoleucine Feedback
Insensitive Threonine Dehydratase/Deaminase of Arabidopsis thaliana
                                    VELFYDVLSPYSWLGFEILCRYQNIWNINLQLRPSLITGIMKDSGNKPPGL--LPRKGLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29; Mismatches 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               line GM11b (omr1/omr1).";
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF17212; AAD54324.1; -
HSSP; P04968; ITDJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HNTERPRO; IPRODO524; -
INTERPRO; IPRODO534; -
INTERPRO; IPRODO134; -
INTERPRO; IPRODO1926; -
PFAM; PF00291; PALP; 1.
PFAM; PF00285; Thr_dehydrat_C; 2.
PROSITE; PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
SEQUENCE 592 AA; 64562 MW; 1D02F73AD53FAF6C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 89; DB 10;
Pred. No. 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         204 ---GATVVLFGDSYDQAQAHAKIRAEEEGLTFIPP 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      592 AA
                                                                                                                                                                                                                                                                                                                                                                                                                FGLPITVAHVDGQTHMLFGSDRMELLAHLLG 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, Crea Olamay-2000 (TrEMBLrel. 13, Last Ol-CT-2000 (TrEMBLrel. 15, Last THREONINE DEHYDRATASE/DEAMINASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.5%; 22.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142 AEKAGMSAEQAQG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09SPF1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      09SPF1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 14
09SPF1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   105
                                                                                                                                                                                                                                                                                                                                                                                                                                              166
                                                                                                                             99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DDT THE READ OF THE PARTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                             ð
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
11;
                                                                                                                                                                                                                                                        STRAIN-CV. COLUMBIA;
Mourad G., Emerlock R., Marion A., Smith A.;
"Cloning and Sequencing of a cDNA Encoding Threonine
Dehydratase/Deaminase of Arabidopsis thaliana (Accession No. AF096281)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----AVNLEHP-EMLEKASRELWMRVWSRNEDITE-------PQSILAA 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AEKAGMSAEQAQG-------LLEKIATPKVKNQLKETTEAACRYGAFGLPITVA 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 PPGL-LPRKGLYMANDLKLLRHHLQIPIHFPKDFLSVM-----LEKGSLS-AMRFLT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52 PPKLPLPR-----LKVSPNSLQ----YPAGYLGAVPERTNEAENGSIAEAMEYLTNI 99
                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-CV. COLUMBIA;
Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
Bowman C.L., White O., Nierman W.C., Fraser C.M.;
"Arabidopsis thaliana chromosome III BAC T22K18 genomic sequence.";
Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                  Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 LSTKVYDIAIESPLQLAKKLSKRLGVRMYLKREDLQPVFSFKLRGAYNMMVKLPADQLAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            160 GVICSSAGNHAQGVALSASKLGCTAVIVMPVTTPEIKWQAVENL--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. Mourad G.S., Smith A.M.; Molecular characterization of the genomic clone, including the promoter sequences, of threonine dehydratase/deaminase from Arabidopsis thaliana.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 592;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (JAN-2200) to the EMBL/GenBank/DDBJ databases EMBL; AF096281; AAC97936.1; -. EMBL; AC010927; AAF04418.1; -. EMBL; AF221984; AAF32370.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           592 AA; 64634 MW; 16658747052FAE7C CRC64;
                                               01-OCT-2000 (TEMBLrel. 10, Last sequence update)
THREONING DEHYDRATASE/DEAMINASE (EC 4.2.1.16).
OMRI OR T22X18.12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.5%; Score 89; DB 10;
22.8%; Pred. No. 13;
tive 29; Mismatches 57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PFAM; PF00291; PALP; 1.
PFAM; PF00291; Thr_dehydrat_C; 2.
PROSSITE; PS00043; HTH GNTR FAMILY; UNKNOWN_1.
PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                189 HVDGQTHMLFGSDRMELLAHL---LGEKWMGPIPP 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204 ---GATVVLFGDSYDQAQAHAKIRAEEEGLTFIPP 235
              592 AA
                                                                                                                                       Arabidopsis thallana (Mouse-ear cress).
                PRT;
                                                                                                                                                                                                                                                                                                                                  (PGR 98-199).";
Plant Physiol. 118:1534-1534(1998).
                                O9ZSS6;
01-MAY-1999 (TrEMBLrel. 10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 7.5%
Best Local Similarity 22.8%
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HSSP; P04968; ITDJ.
INTERPRO; IPR000524; -.
INTERPRO; IPR000634; -.
INTERPRO; IPR001051; -.
INTERPRO; IPR001926; -.
                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                         NCBI_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                98-199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                988Z60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lyase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         105
9SSZ60
                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Db
```

Search completed: April 3, 2001, 11:39:53 Job time: 93 sec

				10		
			2			
					è	·
1.						
1.6						
				•		
					•	
	¥.		· · · · · · · · · · · · · · · · · · ·			
					**	
		F)				
			·			:
•						•